IDA

INSTITUTE FOR DEFENSE ANALYSES

A Multimedia-Based System for Foreign Language Listening and Reading Comprehension

Dale E. Lichtblau

19990624 022

Preface

This report presents a conceptual design for a multimedia-based system for foreign language listening and reading comprehension. It describes how multimedia, artificial intelligence technology, and the World Wide Web may be combined to provide an innovative and effective approach to foreign language learning in the Department of Defense. The work was performed as part of a Central Research Project titled *Multimedia Studies*.

Discussions of many of the ideas advanced in this paper were held with Dr. Jared Bernstein (Entropic Research Laboratory), Dr. Philippe Delcloque (Napier University, Scotland), Dr. Nina Garrett (The Center for the Advancement of Language Learning), and colleagues from the Institute for Defense Analyses: Dr. Brian A. Haugh, Dr. Judy Popelas, and Dr. Craig Will. For their valuable insights, special thanks are in order.

Dr. Ken Reed, Curriculum Specialist for Foreign Language and Technology for the Alexandria Public School System in Virginia, reviewed the various discussions of German grammar presented throughout the paper. He also graciously checked my English translations of the various German examples. His help is greatly appreciated.

The following Institute for Defense Analyses research staff members were reviewers of this document: Dr. Dexter Fletcher, Dr. Brian A. Haugh, Dr. Richard J. Ivanetich, Dr. Eric W. Johnson, Dr. Michael R. Kappel, Dr. Richard P. Morton, and Ms. Christine Youngblut. The author gratefully acknowledges their contributions.

Finally, I wish to thank Katydean Price, Senior Technical Editor, Institute for Defense Analyses, for her inimitable editorial assistance.

Table of Contents

List of Figures and Tables

Executive Summary

Foreign language learning software doesn't have to be boring and ineffective. We can build a compelling and effective language learning system—one firmly grounded in second language acquisition theory—using existing and proven technology: multimedia, artificial intelligence, and the World Wide Web. This report describes one approach.

Benefits to the Department of Defense

The Department of Defense's need for skilled linguists continues to grow. Its Defense Language Institute in Monterey, California, is the world's largest. There is always a need for good language learning tools in Monterey and throughout DoD. The major results of my Central Research Project are a functional description and conceptual system architecture of an innovative language learning tool. The tool should help DoD meet its foreign language needs.

Overview of Proposed System

The Web, with its wealth of authentic foreign language material, is a tremendous but largely untapped resource for foreign language study. The key to unlocking this resource is a system that can reformat Web-based foreign language material to promote comprehension. Comprehension leads, in turn, to second language acquisition. I propose a system that uses common typographical and not-so-common speech processing techniques to improve foreign language reading and listening comprehension. Because comprehension is recognized as fundamental to language learning, the system should enhance foreign language learning. Using typographical devices such as **bold-face type**, color, and underlining, the system will emphasize the grammatical *cues* that convey meaning and advance the overall understanding of the written language. Similarly, using speech manipulation techniques to modify articulation, loudness, pitch, and timing, the system will help clarify for the listener the grammatical structure of the spoken word.

Natural language processing (NLP) and speech processing (SP) are the crucial technologies that make this possible. NLP provides the linguistic analysis and subsequent syntactic

encoding of Web-based foreign language material. The encoded written material can then be presented—using common text formatting tools—to the language learner in a way that reveals and clarifies the grammatical structure of the foreign language. SP technology will enable the modulated playback of spoken material in a way not unlike that used for text formatting: it can emphasize subject phrases; it can lengthen inter-clause pauses; and it can sharpen the articulation of individual words. It will enhance listening comprehension and, ultimately, language learning.

Example

Imagine that you are browsing an interesting foreign language Web-site. You're trying to learn the foreign language and have an avid interest in the topic of the site. A browser "helper program" automatically reformats the page, simply highlighting each subject and verb clause in each sentence. With these small cues, you now easily unravel the syntax and grasp the main idea. Pictured is just such a reformatted snippet from the Web site of the on-line version of the German newspaper *Die Welt*.[2]

dpa Bonn - ***Die 16 Bundesländer haben*** endgültig grünes Licht für die Reform der deutschen Rechtschreibung ***gegeben***. Bis Ablauf des offiziellen Einspruchstermins, gestern 12.00 Uhr, wurde kein Veto mehr eingelegt, ***teilte die zuständige Staatskanzlei Schleswig-Holsteins mit***.

© DIE WELT, 6.3.1996

Subject Verb

Underlying Assumptions and Design Objectives

Unlike most of the foreign language learning software available today, the system I am proposing is firmly grounded in second language acquisition theory. Four complementary assumptions about critical aspects of the language learning process led to the system's overall design:

1. Krashen's *Comprehensible Input Hypothesis.* According to Steven Krashen (1982), input comprehension is fundamental to the language acquisition process. But comprehension is not limited to our linguistic competence; it also depends on context, our knowledge of the world, and other extra-linguistic information. Exposure to lots of authentic, understandable foreign language material is absolutely vital to successful language learning. The Web will provide the authentic material; the proposed system will provide the environment to make that material more comprehensible.

2. *MacWhinney's Competition Model.* Elizabeth Bates and Brain MacWhinney (Bates 1987, MacWhinney 1987, MacWhinney 1992) have argued that regardless of the value

---

[2] An English translation of this and the other German examples used in the paper can be found in Appendix D.

of extra-linguistic information, linguistic-based comprehension is fundamentally a process of linguistic cue acquisition. Linguistic cues are the surface-level phonological and morphological features (e.g., case endings) that map to underlying meaning or interpretation. NLP and SP technology can identify these important cues in the authentic and Web-based material.

3. *Comprehension Facilitated by Cue Accentuation Hypothesis*. I propose—and this is the basic idea underlying the system—that by being sensitized to the important linguistic cues of the target foreign language the student will comprehend more of the input. Typographic and acoustic accentuation of these cues can have, therefore, a marked positive effect on foreign language learning. I propose using commonplace multimedia techniques—typography, graphics, sound—to sensitize the user to these cues.

4. *Oxford's Different Learning Styles Theory*. Finally, Rebecca Oxford (Oxford 1995) has argued that it is vitally important that the learner be able to tailor the computer-aided learning environment to match his or her (possibly idiosyncratic) learning style. Accordingly, the proposed system will give the user maximum possible control over the environment.

Recapitulation

The integration of multimedia and artificial intelligence technology into a Web-centric application can bring the world's languages within easier grasp of the DoD language student. A sound set of second language acquisition assumptions provide a firm basis for the system's design objectives. The proposed system is a practical approach to foreign language learning and can provide genuine value to DoD.

Chapter 1. Introduction

1.1   Background

The original objective of this Central Research Project was to explore the potential for multimedia computer technology in the Department of Defense (DoD). It soon became apparent, however, that the original scope of the effort was too broad. Multimedia had come to encompass a large range of technologies—from simple sound and three-dimensional color graphics to full-motion video and "virtual reality"—a range that was too extensive to allow for anything but a superficial treatment given the available Central Research Project resources and time constraints.

In order to narrow the scope of the research while still remaining true to the original multimedia orientation and the technology's applicability to DoD, the focus of the Central Research Project was adjusted in favor of a search for an application for which multimedia technology could be considered an ideal or natural fit, or which, in the absence of multimedia, could only be imperfectly or partially realized. In other words, the new objective was to identify an important application which would leverage the capabilities, if any, that could be considered to be unique to multimedia technology.

Although the original intent was to focus on multimedia, it also became clear that any such research could not be kept separate from other developments in the world of information technology. The growth of the Internet and its multimedia-rich cousin, the World Wide Web, for instance, could not be ignored. The question then became, could the resources of the Internet be coupled with multimedia technology to create an innovative application that would be of both interest and importance to DoD?

These early considerations led to the idea of using multimedia technology and the Internet for foreign language learning, particularly if speech processing (SP) and natural language processing (NLP) technologies from the field of artificial intelligence (AI) could also be harnessed in some useful way. It seemed that a foreign language learning application might be an

1

excellent application of multimedia technology, especially if coupled with the Internet and language-related AI technologies.

The conceptualization of such a foreign language learning system thus became the objective of my research. This report reports the results of that research. Generally, these results consist of a conceptual specification of a system^2 for foreign language learning. The system is based on multimedia technology, coupled with speech processing and natural language processing technologies. Input to the system is genuine foreign language material in the form of newspaper articles, radio broadcasts, city travel guides, archived material, and other authentic text. Output of the system is a modified form of the input, a form intended to be more easily understandable to the foreign language learner. Specification of the form of the system output is under the user's control. Although German is used for the foreign language examples, the results are applicable to any language that is widely available on the Web. The approach could be used by non-native speakers to learn English. I chose German as the example foreign language because it is the one I happen to know a little about.

1.2 Relevance to DoD

The importance of foreign language skills in DoD is often overlooked, yet DoD is the largest consumer of foreign language skills in the country, if not the world. The Defense Language Institute in Monterey, California, is the world's largest language training facility with about 3,000 students attending the school annually. As DoD's need for skilled linguists grow, the importance of effective and efficient foreign language training also grows. Accordingly, there is a place in DoD for effective computer-based tools to promote the acquisition and maintenance of foreign language skills. If multimedia technology can enhance in a significant way the effectiveness of DoD's extensive foreign language programs, then it deserves careful consideration.

1.3 Project Objectives

This Central Research Project addresses two objectives:

- Show how multimedia technology can be used in an interesting and important way to meet current or future DoD needs.

- Show the advantages of multimedia technology over other forms of computer technology in a specific domain such as foreign language learning. It addresses the

---

^{2} The terms "system," "application," and "tool" are used interchangeably in this paper.

question, what can multimedia offer over other computer technologies in the field of foreign language learning?

Technology is not new to foreign language instruction. The phonograph record and the audio-lingual method were introduced in the mid-1940s. In the early, pioneering days of computer-assisted instruction (CAI), computer technology also appeared in the foreign language classroom. These early computer applications were of the basic "drill and test" variety, and they were boring and largely ineffective. In recent years with the incorporation of colorful graphics and sound, commercial foreign language CAI tools have become much more compelling and arguably more effective. However, they are still basically just a substitute for teacher-directed drills and tests.

The real challenge, as Nina Garrett says for the promotion of multimedia applications in foreign language education,[3] is for them "to do things that cannot be done otherwise, rather than provide an expensive electronic version of other media" (1987, p. 183). I will argue that the multimedia-based system being proposed herein can do things that cannot be done otherwise: the system can accelerate foreign language learning by automatically making genuine foreign language material more comprehensible to the language learner by using multimedia technology and speech and natural language processing techniques. This capability—to enhance comprehensibility[4]—is not something that can be done otherwise, with other media, or with other pedagogic techniques.

1.4   Purpose of the Report

The purpose of this report is threefold:

1. To provide a conceptualization of a multimedia-based system for foreign language learning.

2. To justify by example the belief that the system is novel and innovative, and that it comes close to achieving the general CAI goal of using computer technology in education to do things that cannot be done otherwise or done easily.

3. To describe the functionality and system architecture of the proposed system in detail sufficient to establish its technical feasibility and to provide a foundation for the later development of a proof-of-concept prototype.

---

[3]   Or in education and training in general.

[4]   In a sense to be explained in Section 3.2.1.

1.5   Organization of the Report

Following this introductory chapter, the report is organized as follows:

Chapter 2 provides an overview of the proposed system.

Chapter 3 presents six design objectives and makes explicit four major assumptions upon which claims of the system's potential efficacy are based.

Chapter 4 is a description of the basic functionality of the system.

Chapter 5 sketches the system architecture of the proposed application.

Chapter 6 is a concluding summary.

Appendix A provides some mocked-up examples of output of the envisaged tool.

Appendix B contains several examples of the way in which "cue detection" seems to be important in developing an interpretative facility with a foreign language, using German as the example language.

Appendix C is a brief survey of some commercial multimedia-based foreign language learning products as well as a few products that implement NLP and SP technologies.

Appendix D provides English translations for all of the German language examples used in the report.

List of References provides the bibliographical details for all cited sources.

List of Acronyms offers an expansion of all acronyms or other abbreviations.

Note: Several pages of the report contain material printed in color. These pages are ES-1, ES-2, 28, 29, 30, 54, and A-4 through A-8.

4

Chapter 2. System Overview

2.1 A Brief Overview

The following is typical German:

Doch es sieht ganz so aus wie die Gasse, wo ich mit Rolf gelandet bin, als wir uns verlaufen hatten.[2]

Three of the words in the initial clause (*doch, ganz, so*) are typical German "filler" words, that is, words used for emphasis, or to express a view of the speaker about what is being said, or to add subtle variations of meaning. They are not essential to understanding the basic, simple message being conveyed, in this case that a certain alley looks a lot like the alley the speaker saw before with Rolf. For the non-native, these filler words tend to obscure the basic syntactic structure of the clause (*es sieht aus wie* - it looks like); they can get in the way of effortless and fluent comprehension of the whole sentence. With little loss of meaning, the sentence could just as easily be written:

Es sieht aus wie die Gasse, wo ich mit Rolf gelandet bin, als wir uns verlaufen hatten.

Or, in the interest of word-for-word fidelity with the original:

Doch** es sieht **ganz** so **aus wie die Gasse, wo ich mit Rolf gelandet bin, als wir uns verlaufen hatten.

The latter uses a simple typographic technique (bold face) to intentionally downplay the filler words (*doch, ganz, so*) and to emphasize the importance of the words that bear the most meaning.

This use of typographic devices to facilitate comprehension of foreign language text (and the use of analogous speech modification techniques for spoken material) is the basic idea of the system being proposed here.

---

[2]   Utta Danella (1991, p. 175). See Appendix D for English translations of all German examples that appear in the report.

The system being proposed is a multimedia-based tool to help in foreign language listening and reading comprehension. The key idea is to present authentic foreign language material to the user in a way that improves reading and/or listening comprehension. The basic meaning of the material is not changed in any way. The material is reformatted in a way that increases its level of comprehension by the user.

Intended Users

The intended user of the proposed system is a foreign language student who is somewhat beyond the beginner level. The system could be used either by the intermediate-level language student to improve his or her basic foreign language reading and listening comprehension ability or by an advanced student for reading and listening comprehension skill maintenance.

System Input

System input is intended to be authentic foreign language material accessed via an Internet browser and downloaded to the system. This material could be either textual or audio, or both.[3] While the intention is to take advantage of the large quantity of foreign language material readily available on the Internet,[4] the system could use any source of foreign language material for input.

System Output

System output is to be reformatted or modified input. The reformatting or modification of the input material is intended to facilitate user comprehension. Typographic techniques are to be used to facilitate comprehension of textual input. Speech rate modification and other speech modulation techniques are to be used to promote greater listening comprehension. The concurrent display of a textual transcript of audio material can also facilitate listening comprehension and is to be a capability of the system.

User Control

The extent, focus, and type of reformatting or modification performed on the source input is to be under the control of the user. The user is to determine whether and how much of

---

[3] Video material is becoming increasingly available on the Internet, and one can imagine how the proposed system could be extended to incorporate a video component. For the present, however, video input is not a capability of the envisaged system.

[4] Although it is estimated that 90% or more of the material on the Internet is in English, there still exists copious quantities of foreign language material, especially in western European languages.

the source input is to be reformatted. He or she is also to determine what elements of the input are to be singled out for reformatting or modification. Finally, the user is to determine the type or form of the reformatting or modification, within the limits of the system's capabilities.

Principal System Mechanisms

The system is to rely on computer-based speech processing and natural language processing technology to provide the lexical-, morphological-, and syntactic-level input analysis that will be necessary to deliver the comprehension-enhancing presentation capabilities that are envisaged. The basic elements of the proposed system are illustrated in Figure 1 and are described more fully in subsequent sections of this chapter and in Chapters 4 and 5.

User Control

Input: Internet-based, Authentic Foreign Language Material	• Speech Processing • Natural Language Processing • Text and Audio Reformatting and Display Capabilities	Output: Foreign Language Input Enhanced for User Comprehensibility
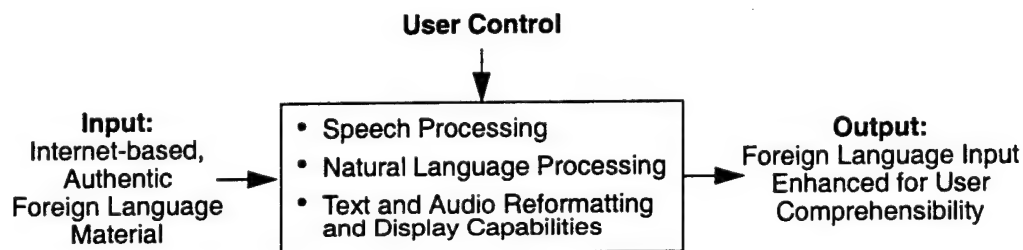
Figure 1. Basic Elements of the Proposed System

2.2 Purpose of the System

The purpose of the proposed system is to facilitate two important aspects of foreign language learning: reading and listening comprehension. It is intended to apply multimedia and certain artificial intelligence technologies to the problem of making authentic foreign language text and audio material more comprehensible to the user.

Reading comprehension will be enhanced by using typographic techniques such as underlining, *italics*, color, font size, **bold face**, simple line graphics, etc., to help clarify the syntactic structure of the textual material being read. Listening comprehension is to be enhanced by using speech processing techniques. These will include both general speech rate modification, and word-, phrase-, and sentence-level modulation of speech articulation, loudness, pitch, and timing to help clarify the syntactic structure of the audio material.

The system is meant to supplement and not supplant normal classroom or other formal tutoring activities for foreign language learning. As a classroom adjunct it could be used to exhibit real world examples of various target language grammatical constructions. There are

7

also research possibilities for the system. These two system aspects are discussed further in Section 4.2.1 and Section 4.2.2.

The proposed system is limited to the two language comprehension skills of listening and reading. It excludes the production skills of speaking and writing.[5] The emphasis on the two comprehension skills is an explicit acknowledgment of what seems to be the special role that input plays in language acquisition.[6] My intent here is that by promoting reading and listening comprehension, the system will be able to facilitate the overall language learning process, including the production ability.

Multimedia technology will play two roles in the envisaged system. First, it will allow simple access to and use of two different foreign language media: textual and audio.[7] Second, it will support the presentation of foreign language material in a comprehensively enhanced way.

Two AI technologies—speech processing and natural language processing—will be used to provide the syntactic analysis necessary to enhance the comprehensibility of the input material prior to presentation using the multimedia capabilities of the system.

One difference between the proposed system and many other multimedia-based systems that are commercially available for foreign language learning is the emphasis on authentic[8] foreign language material. This is in contrast to what is often called the "synthetic" material characteristic of text books, graded-level readers, language learning audio tapes, and so forth.

2.3   Intended Users

Intended users of the proposed system are foreign language students. Because the system is to facilitate comprehensibility by accentuating syntactic-level features that can aid in interpretation, some familiarity with the basic grammar of the target language is to be assumed. The system will be able to be used by the *advanced beginner* to gain further understanding of

---

[5]   Jacques Barzun (*1991*, p. 29) has challenged this traditional *four skills* view of foreign language learning, claiming that there are not really four distinct skills at all, but rather *four modes of one power* [my emphasis], that is, the ability to do something at will.

[6]   This claim is explained in Section 3.2.1.

[7]   As noted earlier, there is a probable future role for full-motion video in the proposed system.

[8]   "Authentic" is used to refer to foreign language material produced mainly by and for native speakers of the language. The term is used in contrast to what is being called "synthetic" material, that is, foreign language material produced primarily with a non-native and generally a language student in mind. Synthetic material if often "graded" in the sense that it is targeted for use by language learners at different levels of foreign language proficiency. Care is usually taken in the preparation of foreign language instructional materials to avoid uncommon idioms and complicated grammatical constructions.

the syntactic features of the target language. The *intermediate-level student* should be able to take fullest advantage of the system to make large gains in reading and listening comprehension ability. The *advanced student* should be able to use the system for comprehension skill maintenance or to work with difficult material from an unfamiliar subdomain, for example, scientific or literary material.

The system could be used in conjunction with a formal language study program or by an individual pursuing a foreign language interest independently.

The reading comprehension feature of the system could be used in the classroom by a teacher to present and explain grammatical features of the target language, using synthetic or authentic input material.

2.4   System Input

The principal input to the system is to be authentic foreign language material downloaded from the Internet. The system would serve as an Internet browser "plug-in" that is invoked to first preprocess and then present foreign language textual or audio material selected by the user.

Foreign language textual material appropriate for language learning could consist of current newspaper or magazine articles; company home pages, including product or service descriptions; city or country travel guides; archived literature and scientific papers; and so forth. The availability of audio language material from foreign language radio and television broadcasters is growing rapidly. These companies are using audio and video servers to enhance of their Internet presence.

Live and recorded audio (language) material is available from hundreds of foreign language radio and television stations.[9] There is a similar number of foreign language newspaper and magazine publishers that offer Internet access.[10]

---

[9]   RealNetworks, Inc. (www.real.com), a leading vendor of streaming audio/video plug-in software for Web browsers, lists over 30 European countries that currently offer foreign language radio broadcasts via the Internet. The list includes: Andorra, Austria, Belgium, Croatia, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Liechtenstein, Lithuania, Norway, Poland, Portugal, Romania, Russian Federation, San Marino, Slovak Republic, Spain, Sweden, Switzerland, Netherlands, Turkey, and United Kingdom.

[10]   The Web site titled "Newspapers of the World on the Internet", for example, recently listed over 240 non-English language Internet-available newspaper from Europe alone. The site is available at http://www.southamerica-business.com/newspapers/europe.html.

While the intention—for reasons to be presented in Section 3.1.2—is to use Internet-based source material for system input, the proposed system is not dependent on the Internet and could use input material from other sources. A teacher in a formal classroom environment, for example, might want to use the system with digitized textbook material as input to present and explain grammatical features of the target language.

2.5   System Output

The output of the proposed system is to be the system input that has been enhanced for comprehensibility by the foreign language learner. Comprehensibility enhancement will result from the accentuation or emphasis of lexical-, morphological-, and/or syntactic-level language cues[11] that are interpretively significant for the target language.

Enhancements are to take the form of typographic modification of the textual material and acoustic modification of the audio material. Typographic modification will consist of underlining, color, font style and size, and similar graphical presentation features. Acoustic modification will consist of speech rate, inter-phrase or sentence pauses, stress, and pitch modification.

2.6   User Control

The user of the proposed system is to be provided the means to control the extent, focus, and type of analysis and reformatting to be performed on the input material. Input material reformatting or modification could range from extensive (requiring a phoneme-level parse) to minor (requiring only phrase- or sentence-level analysis).[12] The focus of the analysis (for example, to identify and then underscore prepositions that govern the dative case) is also to be under user direction. The user is also to have control over the presentation-level devices or techniques to be used to enhance the comprehensibility of the input. For display of text, for example, one user might prefer the use of color to draw his or her attention to the interpretive cue or feature of interest while another user may prefer the use of a different font style or font size.

Higher-level controls, such as whether audio and textual material are to be presented simultaneously or not, are also to be under user control.

---

[11]   Cues are the surface-level (i.e., non-semantic- or discourse-level) features of a language that map to an underlying meaning and that enable a person to interpret or understand what is being said. See Section 3.2.2 for a fuller discussion of surface-level cues.

[12]   It is not being suggesting that phrase- or sentence-level analysis does not require any lower-level *analysis*. Linguistic analysis is largely a "bottom-up" process.

2.7 Principal System Mechanisms

There are three principal system mechanisms required to provide the basic functionality of the proposed system: speech processing technology, natural language processing technology, and multimedia presentation capabilities.

Speech processing technology is to be used for both audio input preprocessing and audio output presentation. Digital audio input preprocessing will consist of either speech alignment processing—detecting and mapping word boundaries in the audio stream to a word-level transcript of the material—or automatic speech recognition (ASR) processing in cases where a transcript is not available. Speech processing technology for audio output presentation will be used for audio playback of the preprocessed audio input and/or speech synthesis of textual material for which an audio counterpart is not available.

Natural language processing is a key component of the proposed system. It is to be used for the lexical-, morphological-, and syntactic-level analysis of the foreign language textual input.[13] The purpose of this analysis is to enable subsequent display of the textual material with user-selected grammatical features emphasized or underscored in the user-specified way.

Multimedia presentation capabilities round out the three principal technologies to be incorporated into the system. Two basic presentation capabilities will be required: text-graphics and sound (speech).

A relatively advanced text-graphics display capability will be required to enable the full range of typographic devices envisaged for the accentuation of the interpretive cues in the textual material. The full set of standard text formatting techniques like font size and type, italics, bold face, underlining, pair kerning, and color will be required. In addition, the use of line graphics (e.g., to indicate the antecedent referent of a pronoun) will be required.

---

[13] NLP may also be used to further analyze and encode output of the ASR engine if sufficient syntactic information is not provided by the ASR engine in conjunction with its digital signal analysis.

Chapter 3. Design Objectives and Key Assumptions

3.1   Design Objectives

The proposed system was conceived with six basic design objectives in mind:

- Multimedia Based - to address the challenge to use multimedia technology to do something that could not otherwise be done in the foreign language classroom.

- Use of Authentic Material - from the conviction that foreign language learning is most successful when the learning process is kept as similar as possible to the process used in native language acquisition and that native language learning is based on authentic input.

- Learner Driven - from the belief that learning, in general, is more successful when the learner can play an active role in the process.

- Use of AI Technology - a necessary condition for the computer-based linguistic analysis that is central to the system.

- Authoring Independent - to enable the broadest possible range of system input material.

- Grounded in Second Language Acquisition (SLA) Theory - to provide a theoretically sound basis for the system.

While these objectives are ultimately arbitrary, they derive from my hypothesis that a computer system that satisfies all of the objectives can be a very effective foreign language learning tool. Each of these design objectives is discussed more fully in the following subsections.

3.1.1   Multimedia Based

The proposed system was originally conceived with the goal of finding a DoD-relevant and important application that would make good use of multimedia technology, that is, the computer integration and management of multiple media.

On the face of it, any cognitive task that can be approached using multiple representational modes (symbolic, pictorial, acoustic) is a good candidate for a multimedia-based solution. The emergence in recent years of multimedia technology[2] now gives the application developer the option to use *integrated* multiple media. Previously, the developer was limited to the use of a single representational mode, chosen from among the multiple forms available, or the non-integrated use of available options. A map application, for example, could now overlay in a tightly integrated way a graphical depiction of a terrain with a symbolic representation, including text and cartographic symbology.

One advantage of a multimedia approach to a cognitive task seems to lie in the fact that multiple media provide an element of mutual reinforcement that is absent from any single medium approach. An illustration or picture can enhance or clarify written material; the text, in turn, can further explicate the meaning of the illustration or picture. Foreign language learning would seem to be a natural application of the inherent synergy of audio and video, in combination with images, graphics, and text.

3.1.2   Use of Authentic Material

The second design objective was to see if there was some way to take advantage of the copious quantities of authentic foreign language material available on the Internet.

The use of Internet-based foreign language material as system input offers several advantages:

- Authentic - The Internet provides a readily available source of authentic foreign language material, that is, material prepared by and primarily for native speakers of the foreign language. The material is not textbook material that has been simplified for pedagogic purposes. It reflects the language as it is spoken and written by contemporary native speakers.

- Parseable - Much of the foreign language material to be found on the Internet is written by professional writers and journalists. As such, it is generally better grammar and more amenable to natural language processing techniques.

- Topical - The Internet provides a large source of information that is topical, that is, of current and "local" interest. This feature is important for a number of reasons.

---

[2]   "Multimedia technology" is used here in the technical sense of computer-based integration and concurrent management for task purposes of *multiple media* such as text, graphics, images, sound, animation, and full-motion video.

Current events are of general interest to people and current news stories convey genuine real-world information. A story in a current newspaper is also one to which a foreign language student is likely to bring a certain amount of background information that will help in comprehension of the foreign language version of the story. The student is also likely to gravitate to foreign language material for which he or she has a personal interest, and this personal interest is likely to provide some language learning motivation. The sheer size of the Internet suggests that the student will be able to find foreign language material about practically any topic of personal interest.

- Available - The Internet is readily available and easily accessible from a personal computer. The proposed system would provide the user with language learning opportunities at any time, day or night.

- Digital - The authentic material available on the Internet is in digital form and can be directly accessed and processed. There is no need to convert input material to a digital form before use.

- Effective - Some recent research by two French language professors, Carol Herron and Irene Seay of Emory University, confirms that listening comprehension improves with increased exposure to authentic speech (Herron and Seay 1998). This same study also suggests that "adjusting levels of speech (speed, content, and form) to student's developing comprehension" may be "helpful to the intermediate-level foreign language student" (1998, p. 2).

3.1.3   Learner Driven

The third design objective for the proposed system is that it is to be learner driven as opposed to a tutorial-based, teacher-driven system. It is to be learner driven in the sense that the user will have primary control over both the grammatical features of the language he or she wishes to have emphasized during presentation and the form or typographic style of that emphasis. The user will be free to explore the complexities of the target language at his or her own pace and in the directions of his or her own choosing. By giving the user primary control, the student will be able to discover the learning strategies and techniques that are most effective in the language learning process.

3.1.4 Use of AI Technology

The fourth design objective was to use AI technology—specifically SP and NLP—in a central way in the system.

There were two reasons for wanting to use AI technology: first, because I believed that having a computer tool readily available that could grammatically parse foreign language material would be of real benefit to the language learner; second, because I felt that relatively *modest* SP and NLP capabilities would be sufficient for a system that was limited to offering listening and reading comprehension support. A language learning system intended to offer full verbal interaction—both understanding and responding to the student's input—would require both robust automatic speech recognition and a high level of semantic- and discourse-level understanding, that is, full NLP. While state-of-the-art NLP is reasonably good at lexical- and syntactic-level language processing, it still leaves much to be desired at semantic- and discourse-level interpretation.[3] Speech recognition of sentence material read aloud in a normal, casual fashion (by native speakers!) still results in errors rates of about 1 in 20 words (Bernstein 1995, p. 23).

3.1.5 Authoring Independent

The fifth design objective is that the system should not require "authoring" of input material. As far as possible, the system is to be indifferent as to the source or nature of its input material.

To understand this objective, one has to look at the considerable authoring effort involved in producing the CD ROM-based multimedia titles that currently characterize much of the foreign language education market. Content preparation easily accounts for the majority of the production costs of multimedia titles and is generally recognized as the one factor inhibiting the growth of the industry. If content development can be more or less automated in real time using the techniques described here, then content development is no longer an inhibiting factor.

Moreover, content authoring would be inconsistent with the goal that the system be learner driven. The system described requires no manual pre-processing of source material.

---

[3] Lexical analysis is processing at the single-word-level; it identifies words and their parts of speech. Grammatical ambiguities (e.g., *covering* used as a present participle of the verb *to cover* or as a singular noun—*the covering*) at the lexical analysis level are resolved at the syntactic analysis level (Smeaton 1992). I include morphological analysis—the breaking down of a word into its sub-word components (typically a base form and a suffix)—as part of lexical processing. Even at the lexical analysis-level, there is still room for improvement. State-of-the-art part-of-speech-taggers operate at only 95% correctness (Levin and Evans 1995).

The system is to be under complete student control with the user selecting the material that the system automatically preprocesses for subsequent viewing and listening.

3.1.6   Grounded in SLA Theory

Finally, the system should be grounded in some sense on the precepts of a reasonable second language acquisition theory. The proposed system is based on Steven Krashen's Comprehensible Input Hypothesis, Brian MacWhinney and Elizabeth Bate's Competition Model, my own Cue Accentuation Facilitates Comprehensibility Hypothesis, and Rebecca Oxford's Multiple Learning Styles Theory.

1. The overall emphasis of my proposed system on reading and listening comprehension is due to Krashen's Comprehensible Input Hypothesis which says, in effect, that successful language acquisition is the result of being exposed to copious quantities of comprehensible input.

2. MacWhinney and Bates takes this hypothesis a step further and argue that language acquisition is a process of cue acquisition, where cues are the surface-level phonological or morphological features of a language that map to the meaning being conveyed.

3. The central premise of the proposed system is based on these two theories: the accentuation of the input cues that are important to comprehension can enhance the comprehensibility of the input for the user and can lead to the successful acquisition of the foreign language.

4. Finally, the emphasis on student control of the proposed system is consistent with the theory of multiple learning styles, a theory recently linked to intelligent computer-assisted language learning by Rebecca Oxford.

Each of these four assumption is discussed more fully in the next section.

3.2   Key Assumptions

The ultimately efficacy of the proposed system will depend on the validity of four key assumptions regarding the nature of foreign language learning, language itself, and the general learning process.

3.2.1   Krashen's Comprehensible Input Hypothesis

The proposed system assumes the validity of Krashen's well-known but not uncontroversial Comprehensible Input Hypothesis (Krashen 1982).[4] Its emphasis on the importance

and special role of comprehensible input in successful second language (L2) acquisition is one of the four key ideas central to the proposed system.

Krashen's Input Hypothesis is actually one of five hypotheses that collectively compose his SLA theory.[5] While these five hypotheses are not unrelated, only his "Natural Order Hypothesis" is necessary to understand the Input Hypothesis.

The Natural Order Hypothesis posits that "the acquisition of grammatical structure proceeds in a predictable order" (Krashen 1982, p.12). For example, acquisition[6] by children and adults learning English as a second language of the progressive marker *ing* as in "He is playing baseball" and the plural marker /s/ generally precedes the acquisition of the progressive auxiliary as in "he *is* going" and the articles *a* and *the*. The important point here is that language acquisition is a process of moving from one level or stage of understanding or comprehension to another in a fairly predictable fashion, with each stage or level characterized by some set or other of grammatical structures. The language acquisition process is the natural process of moving from stage i, where i represents current competence, to stage $i+1$, where $i+1$ represents the next order of competence in the target language. (It should be noted while the Natural Order Hypothesis is called a hypothesis, there is actually considerable evidence for its general correctness.)

The Input Hypothesis claims that "a necessary (but not sufficient) condition to move from stage i to stage $i+1$ is that the acquirer understand input that contains $i+1$ *[structures]*..." (Krashen 1982, p. 21). In other words, the language learner (or "acquirer" in Krashen's technical sense) progresses through the natural order of language acquisition by being presented with and in some way comprehending structures that are beyond the learner's present competence.

But how is this possible? How can the learner comprehend structures that, by definition, are beyond the learner's present competence? The answer, according to Krashen, is that we appeal to more than just our linguistic competence.

---

[4] Krashen's Comprehensible Input Hypothesis has been criticized recently as being "too abstract to be tested" (Oxford 1995), yet other researchers consider Krashen's Input Hypothesis "to be one of only a handful of consistently supported research findings" (Hubbard 1995).

[5] The other four are the [Language] Acquisition-Learning Distinction, the Natural Order Hypothesis, the Monitor Hypothesis, and the Affective Filter Hypothesis.

[6] Language *acquisition*, according to Krashen's first hypothesis, is to be distinguished from language *learning*. Language acquisition is "a process similar, if not identical, to the way children develop ability in their first language." Language learning, on the other hand, refers to "conscious knowledge of a second language, knowing the rules, being aware of them, and being able to talk about them" (1982, p.10).

18

[W]e use more than our linguistic competence to help us understand. We also use context, our knowledge of the world, our extra-linguistic information to help us understand language directed at us. [Krashen 1982, p. 21]

The Input Hypothesis further posits that if enough input is available and this input is understood or comprehended, then *i+1* acquisition will occur pretty much automatically.

It should be clear how several of the proposed system's design objectives can be justified by Krashen's Input Hypothesis. The use of authentic material will assure the availability of sufficient quantities of *i+1* input. The use of Internet-based material, selectable by the user, will ensure that the user will be bringing to the comprehension task some real-world background knowledge and contextual information sufficient to understand the messages being conveyed. Extra-linguistic information will be provided by multimedia-based graphics and images that may accompany the linguistic material.

The crucial point regarding Krashen's Input Hypothesis, however, is the assumption that *input comprehension is fundamental* to the language acquisition process. Language learning proceeds apace with language understanding. The multimedia-based system being proposed here is intended to facilitate that understanding and thus facilitate foreign language learning.

3.2.2 MacWhinney's Competition Model

Brian MacWhinney and Elizabeth Bate's Competition Model (Bates and MacWhinney 1987; MacWhinney 1987, 1992) provides another perspective of the language acquisition process. The validity of this perspective is also a key assumption of the proposed language learning system.

While Krashen downplays the role that strictly linguistic information plays in the process of moving from stage *i* to stage *i+1* in language acquisition (appealing instead to a special, often overlooked role for context and extra-linguistic information), MacWhinney has argued that language acquisition is fundamentally a process of linguistic *cue* acquisition. Whatever else it may be, the acquisition of a language, whether the first or a second language, is at bottom a process of acquiring the surface-level cues that map to function and convey meaning. Linguistic forms may not be the entire language learning story, but they are a fundamental character in that story.

Cues are the surface-level phonological or morphological features or forms that map to an underlying meaning or interpretation. They include preverbal position (subject-verb-object, or SVO) or word order in general, case-markings, subject-verb agreement, animacy, and so

forth. For example, in English, preverbal position is a very strong and highly reliable sentence cue to the function of "subject" or "agent." These cues "compete" (hence the name, "Competition Model") with each other for relative importance or primacy. They compete with each other within the same language. For example, in English, the preverbal positioning cue competes with subject-verb agreement and animacy cues, with preverbal positioning usually winning out. During the language acquisition process, the child learns the relative importance of these cues and thus acquires the native language by ensuring consistency between the emerging cognitive system (the language) and his or her environment.

Cues can also be a factor in L2 learning, particularly if the relative strength of the cues differ between the learner's first language and the target language. As noted above, in English, preverbal position is a very strong cue to the subject function. In German, however, it is usually the case-marking cues of articles and personal pronouns that are critical for the correct interpretation. (Appendix B describes some of the case marker cues widely used in German.) The native English speaker learning German will find preverbal position cues competing with case-marking cues. The German speaker might easily confuse the first noun of an English NNV (noun-noun-verb) sentence with the subject given the ease with which SOV (subject-object-verb) sentences can and frequently are constructed in German using case-markers. The native English speaker's sensitivity to the preverbal position cue is used to identify the subject of an English sentence, while the German speaker is more attuned to case-markers. These habits have to be unlearned when approaching a foreign language. This last remark leads to the fundamental idea and assumption of the proposed system.

3.2.3   Comprehensibility Facilitated by Cue Accentuation Hypothesis

Krashen's Input Hypothesis says that L2 acquisition is a process of being exposed to lots of *comprehensible input*. MacWhinney's Competition Model says that language acquisition is a process of *cue acquisition*, where cues represent the link between sound waves or ink marks on a page and the *meaning* conveyed by those sound waves and ink marks. Language acquisition emerges from comprehensible input, and comprehensibility is a function of linguistic, surface-level cues.

Going one step beyond these two key assumptions, I am hypothesizing that

foreign language input comprehensibility can be facilitated by sensitizing the student to the important linguistic cues of the foreign language.

It is further assumed that

typographic or acoustic accentuation of the cues that play an important seman-
tic role in the target language can have a marked positive effect on input com-
prehensibility and, hence, on foreign language learning.

Cue accentuation may be especially useful if those cues are not used or are used in a confounding way in the learner's native language

It is assumed, for example, that if the language learner can be sensitized to the importance of case markers in a strongly case-based language such as German by highlighting or in some other way emphasizing the German case marking system, then the German input will be more comprehensible and will facilitate the language acquisition process. This is key basic premise of the proposed system.

3.2.4 Oxford's Multiple Learning Styles Theory

The forth and final assumption that has guided the system concept is that it is important to tailor computer-based learning environments to match the often idiosyncratic learning style, proficiency level, and interests of the individual language learner.

Following Rebecca Oxford (1995), it is assumed that a language learner will be more successful if the learner can easily tailor the learning environment to match the learner's own (perhaps idiosyncratic) learning style and thereby consciously reflect on the linguistic patterns of the target language. Accordingly, the proposed system gives the user a variety of controls and tools to create an individualized language learning and language exploratory environment. The user can focus on listening comprehension tasks with or without an accompanying transcript. Alternatively, the student can concentrate exclusively on reading comprehension, using the system's text formatting tools to clarify linguistic features of the material. The set of structural analysis tools is designed to allow the user to also design the form of the on-screen typographic devices used to convey information.

Chapter 4. System Functionality

The proposed system is aimed at the needs of the individual language learner. Most of this chapter describes the envisaged functionality of the system from the perspective of the individual user. The system could just as easily be used in a classroom and could capture data of interest to the SLA research community.

4.1   Individual Use

For the individual user, the proposed system is intended to function as an Internet browser "plug-in" with content selection determined by the user when browsing the Web.[2] The user uses the navigational facilities of the browser to find and initially download foreign language material of interest.

After finding interesting source material, the user invokes the tool as a comprehension aid. If the material is textual, the application is to provide reading comprehension support. If audio material audio is available, the application is to provide support for listening comprehension. In some instances, both forms will be available concurrently, and the user will be able to listen to the audio version while following along with a written transcript. A set of basic controls is to be provided to allow the user to control both the listening and reading comprehension support functions concurrently or to toggle between the two modes.

4.1.1   Reading Comprehension

For textual material, text-formatting and graphics capabilities are to be employed with the syntactic information gleaned from linguistic pre-processing to sensitize the user to surface-level features, the understanding of which is important to fluent reading comprehension. The basic functional controls designed to support reading comprehension are

---

[2]   The proposed system would not be limited to the Web. Its basic functionality is applicable to any of the information resources of the Internet, including anonymous FTP (File Transport Protocol) sites, Archie servers, the Wide Area Information Server, Gopherspace, Veronica servers, and Newsnet.

described in the following subsections and summarized in Table Table 1, "System Functionality for Reading Comprehension Support," on page 25.

Two basic sets of presentation controls for written material are to be available: paging and format. Paging controls are to provide the basic means to navigate through textual material. Format controls are to enable the user to select the surface-level feature(s) for which reformatting support is desired and also to specify the form that reformatting is to take. The purpose of this reformatting is to enhance comprehensibility of the input by making the material in a sense more "transparent" to the user.

4.1.1.1 Paging Controls

Paging controls are intended to allow the user to control both the amount and the pace of the presentation, giving the user the ability to navigate through the material at a pace with which he or she is most comfortable. Paging controls are to be provided in addition to the usual scroll-bars that can be used to scroll the current screen or window.

Written material is to be presented a page (i.e., full-screen or window), paragraph, sentence, or phrase at a time, with the number of paragraphs, sentences, or phrases displayed during each "new page" cycle settable by the user.

The display of a new page, paragraph, sentence, or phrase is to be either "on demand" by the student or under automatic control of the system.

On-demand page turning is to be effected by the user clicking on a "next" button, resulting in the display of the next page, paragraph, sentence, or phrase of the text. In a one-paragraph-at-a-time mode, for instance, the initial paragraph would be displayed at the top of the screen, with each subsequent paragraph being displayed beneath each just-displayed paragraph, and automatic scrolling off the top of the screen of paragraph-size units of material. On-demand page turning would work in conjunction with normal scroll-bar features, allowing the user to scroll, within the displayed material, a display line or user-selected unit (e.g., page, paragraph, sentence, phrase) at a time.

An automatic page turning feature is also to be provided. The user will be able to set a rate (e.g., every 90 seconds) at which each successive page, paragraph, sentence, or phrase of text will be displayed automatically by the system. Automatic page turning can be useful in developing reading comprehension fluency by forcing the student to read quickly for main idea rather than for word-for-word interpretation. It can also relieve the user of the need to manually step through material using the on-demand feature.

Table 1. System Functionality for Reading Comprehension Support

Major Function	Sub-Function	Mode	Description
Paging	Page		New text is displayed a page (screen) at a time.
	Paragraph [count]		New text is displayed [count] paragraphs at a time.
	Sentence [count]		New text is displayed [count] sentences at a time.
	Phrase [count]		New text is displayed [count] phrases at a time.
	Each of the two paging modes is applicable to each of the four paging sub-functions.	Automatic [interval]	New text is presented at regular, user-specified [interval].
		On Demand	New text is presented on user-request.
Format	Page Layout		Paragraph, sentence, phrase, word spacing; hyphenation suppression.
	Morphology		Emphasis of case markers, prefixes, suffixes, endings, etc.
	Syntax		Typographic-based indication of subject-, predicate-, prepositional-phrases; clauses; pronouns, etc.
	Semantics		Proper name flagging, dictionary lookup, vocabulary in context, idiom flagging, register detection.
	Text-to-Speech		Conversion of digitized text to speech.
	Each of the three format modes is applicable to each of the five format sub-functions.	Typography	Font type, size, and weight (bold); color; shading; character and line spacing; simple line graphics
		Always	User-specified cues are always accentuated.
		On Demand	User-specified cues are accentuated on user demand. User selects page, paragraph, sentence, or phrase for which cue accentuation is being requested.

25

4.1.1.2 Format Controls

Format controls are to provide the user with control over the physical display of the input material. The user will specify which surface-level[3] features of the material are to be reformatted and in what form reformatting is to take. This will result in a system-internal presentation specification.

The user will be able to toggle the presentation specification between "always" and "on demand." As the term implies, "always" means the user-specified format is always invoked during the display of the input material.

"On demand," in contrast, means that the presentation specification is to be applied to a user-selected subset of the already-represented material. On-demand formatting would first require the user to select a page, paragraph, sentence, or phase for which reformatting was desired. The user would then click on a "reformat" button that would result in the system re-displaying the selected passage according to the predefined presentation specification already in effect or according to a new ad hoc specification to be provided by the user.[4]

The user will be able to control five different facets of the output display: page layout, morphology, syntax, semantics, and text-to-speech. Each of these is explained in the next subsections.

4.1.1.2.1 Page Layout

Page layout refers to the physical layout and display of the textural material on the display device and includes word, phase, sentence (line), and paragraph spacing; font type, weight, color, and size; and hyphenation. The proposed system is to provide the ability to control all of these layout dimensions of the display.

For example, simply making the phrase structure of a sentence more obvious can be of help in visual parsing and comprehension. This can be done by ensuring that there is a slightly larger and fixed space between each major phrase and larger than usual phrase punctuation.

A simple increase in line-spacing can also provide helpful "white space" to a language learner intimidated by a page literally too full of a foreign language. Larger and simpler (e.g.,

---

[3] The term "surface-level" is being used here to refer to both non-syntactic typographic features (e.g., line-spacing, hyphenation) of the foreign language material and syntactic features (e.g., case markers, subject-verb agreement).

[4] On-demand format control has a potentially important system performance advantage: input material does not have to be pre-processed in its entirety.

sans serif) type faces can also be used to make a foreign language more approachable and less "opaque." The technique is common in the primers used for beginning native language reading instruction.

Hyphenation is another layout artifact that can cause unnecessary confusion among language learners, particularly if the hyphenation conventions of the target language differs from those of the student's native language.[5] The system will allow hyphenation to be suppressed altogether or automatically converted, if possible, to approximate the conventions of the learner's native language.

Figure 2 illustrates these layout techniques for enhancing reading comprehensibility. The same German sentence is rendered in two noticeably different formats. The first uses a default word spacing and line justification algorithm. The second avoids hyphenation, uses a larger than usual space of fixed size between each major phrase in conjunction with larger than usual phrase punctuation (commas and period), a san serif type face (Helvetica), and 14-point as opposed to 12-point line spacing.

Dann müßte also auch der Durchgang, in dem die Tür war, aus der Pia damals heraustrat, ganz in der Nähe sein.

Dann müßte also auch der Durchgang, in dem die Tür war, aus der Pia damals heraustrat, ganz in der Nähe sein.

Figure 2. Page Layout Reformatting Example

4.1.1.2.2 Morphology

Morphology is the study of word-forming elements (morphemes) and processes in a language. These elements (e.g., endings *-ed*, *-s*, *-ing* in English) are the smallest linguistic unit of meaning. Understanding of these elements and the role they play in a language is critical to the learning of that language. The proposed system is intended to promote the acquisition of these elements by sensitizing the user to the important role they play in the language. Common formatting devices are to be used to accentuate or highlight various morphological elements found in the input material.

There are many morphological elements that play important roles in any given language. They also differ, of course, by language. Two examples—case markers and prefix compounds—will illustrate the point.

---

[5]   Even native German speakers are known to sometimes stumble over hyphenated German words.

Case-markers provide a rich source of interpretative cues in many languages, especially highly inflected Indo-European languages. In strongly case-based languages (such as German), the relationship between the verb and the other parts of the sentence is determined mainly by the cases of the constituent parts. Different cases are indicated by case-markers (usually word endings) that serve as important cues to the correct interpretation of the sentence. In the following German sentence, for example, there are five individual case markers (in bold face), denoting two distinct cases (nominative, dative):

> Dann müßte also auch **der** Durchgang, in **dem die** Tür war, aus **der** Pia damals heraustrat, ganz in **der** Nähe sein. [Dannella 1991, p. 175]

The importance of these case-markers is sometimes underappreciated by the language learner, particularly if his or her native language is one in which the case system is no longer entrenched (e.g., English). The proposed system is intended to sensitize the user to the importance of case-markers in a strongly case-based language by employing accentuation techniques to call the user's attention to these morphological cues. The two different cases denoted by the definite article endings in the above quote could be made explicit using different colors (blue for nominative, red for dative):

> Dann müßte also auch **der** Durchgang, in **dem die** Tür war, aus **der** Pia damals heraustrat, ganz in **der** Nähe sein.

Other typographic techniques (e.g., underlining, different font sizes or types) could be used to draw the same distinctions.

Languages use prefixes to form new words from root words. The *ent* prefix, for example, is common in German and is used to change in a fairly predictable way the meaning of the word to which it is prefixed. For example, *hüllen* means *to wrap* or *to cover*, whereas *enthüllen* means *to unwrap* or *to uncover*. The prefix *ent* functions to negate or reverse the meaning of the term it precedes. Native Germans see *ent* as a morphological unit.

However, this morphological unit may compete—in the Competition Model sense— with the common English prefix *en* when immediately followed by a *t* in words like *entomb* and *entitle*. The native German speaker learning English has to learn to see *entomb* not as the vaguely German-like *ent-omb* but as the English (really French) *en-tomb*, and to recognize *en* as an important English prefix.

Typographic techniques can be used to lessen some of the confusing effects of these and similar competitions. Unobtrusive "thin spaces" can be used to "break" competing native lan-

28

guage letter groups (e.g., *entomb* vs. *entomb*) or a conspicuous vertical bar could be used to emphasize the importance of *en* as an English prefix (*en|tomb*) in contrast to the German *ent*.

Novel typographic spacing techniques might also facilitate word meaning guessing—and, hence, comprehension—by making explicit the simple words from which intimidating compounds are formed. Native speakers are often oblivious to the obvious way in which many commonplace compounds are formed: air-port, sky-scraper, seat-belt. We learn these words as "opaque chunks,"[6] seldom aware of their seemingly transparent etymology.

The German *Geschwindigkeitbegrenzung* (speed limit) serves as a good example. It comes from *Geschwindigkeit* (speed) and *Begrenzung* (limit). *Geschwindigkeit*, itself, is built up from the adjective *geschwind* (quick, swift), and the common suffixes, *-ig* (swift-y), and *-keit* (swifty-ness). Similarly, *Begrenzung* is formed from *Grenze* (border, limit), *begrenzen* (to limit), and the common noun-forming suffix *-ung*. It is not hard to envisage the proposed system, using the results of this morphological analysis of the word *Geschwindigkeitbegrenzung* to display the word something like:

Geschwindigkeits|begrenzeung

where the red vertical bar is used to break the compound into its two principal parts, and where blue is used to highlight the two key root words, *geschwind* and *Grenze*.

4.1.1.2.3 Syntax

Syntax refers to the system of rules of a language that govern the way words can be put together to form meaningful phrases (noun, verbal, prepositional), clauses (subject, predicate, dependent, independent), and sentences. This syntactic system includes rules for subject-verb agreement in number, person, and gender; prepositions that govern various cases; verb tense, aspect, voice, and mood; coordinating and subordinating conjunctions; and personal, relative, and demonstrative pronouns.

To quickly recognize the syntactic units of a foreign language and to understand the role they play in conveying meaning in that language are important stages of the second language acquisition process. The proposed system is to be able to help the language learner at these vital stages of the learning process.

The proposed system will allow the individual user to identify the general syntactic element or elements with which he or she is interested. The system will then display the input

---

[6] The expression "opaque chunks" is Douglas R. Hofstadter's. See his recent book (1997, p. 496 ff.) for an interesting discussion of the imagery of compounds surprisingly hidden from native speakers.

material with the selected syntactic elements highlighted or emphasized in the way specified by the user in the user-defined presentation specification.

The highlighting techniques are those that have already mentioned: font style and weight, color, spacing, and underlining. Graphics can also be used in imaginative ways, for instance, to indicate the antecedent references of personal or relative pronouns.

Figure 3 illustrates the use of color and line graphics to highlight clausal subject-verb pairs and relative pronoun antecedents. In the figure, the color blue is used to indicate the subject-verb pairs within each clause; arrows are used to map relative pronouns to their antecedents.

Doch es sieht ganz so aus wie die Gasse, wo ich mit Rolf gelandet bin.

Dann müßte also auch der Durchgang, in dem die Tür war, aus der Pia dam-

als heraustrat, ganz in der Nähe sein.
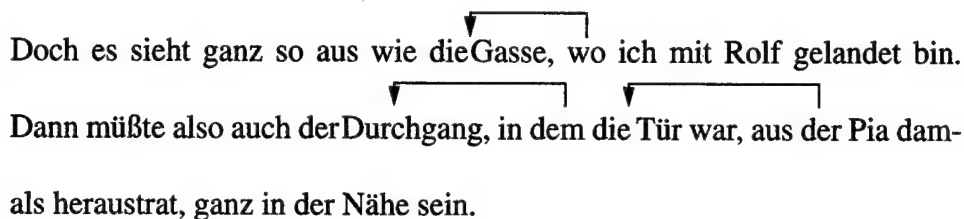
Figure 3. Syntactic Element Highlighting Example

4.1.1.2.4 Semantics

Semantics refers to the specification of the *meaning* of morphemes (the smallest meaningful parts of a word), words, phrases, or sentences. In general, the purpose of computerized semantic analysis of language is "machine understanding" of the material being analyzed. The ability to understand what a newspaper story is about, in order, for example, to alert a human to the existence of the story is a typical goal of the machine understanding community.

Computer-based semantic analysis of foreign language material, on the other hand, is usually expected to result in machine translation (MT) of the source to some preferred target language.

Extensive use of the semantic analysis capabilities of NLP—including MT—is not being proposed for the system being described here. Successful semantic analysis is probably an order of magnitude more difficult than syntactic analysis, while syntactic analysis is, likewise, probably an order of magnitude harder than lexical analysis.

While MT is not to be offered as a functional feature of the proposed system, some modest semantic-based capabilities are being considered. These include proper noun flagging, dictionary look-up, vocabulary-in-context, idiom flagging, and register detection.

30

A simple but potentially useful device for rendering foreign language text more understandable is to highlight those elements that do not, strictly speaking, convey meaning at all, that is, proper nouns or names. One would think it easy to distinguish normal language elements from names of persons, geographical places, institutions, and so forth in a foreign language. It's not. (The capitalization of all nouns in German compounds the problem for the German language student.) Any foreign language learner will attest to the countless times they've wasted time trying to lookup an unfamiliar proper name in a dual-language dictionary before realizing they were dealing with a person's name or some other non-connotative designation. These elements can be highlighted, signaling to the learner that the element is functioning to refer to something and will have little other semantic import.

Although the proposed system is not intended to support foreign language translation,[7] the ability to quickly and effortlessly look-up the meaning of an unfamiliar word in an automated dual language dictionary is likely to facilitate comprehension and promote second language acquisition.

A vocabulary-in-context feature would highlight all source material occurrences of the entries on a student-maintained "stop list." If the student-selected words are common enough and yet rich in multiple meanings, the student should be able to better understand how they vary their meaning when seen in many different contexts.

Somewhat similar to the proposed vocabulary-in-context feature, the system should be able to flag (i.e., highlight) the more common target language idioms. The meaning of these expressions cannot be deduced from their constituent parts but can often be guessed at from the context. Dual language dictionaries often include a large number of common idiomatic expressions.

One problem that arises when using authentic material for language learning is the presence of *register variation*. Register variation is the variation in language forms used by native speakers. These different forms are determined by subject matter, situation, and medium, and usually manifest themselves as different levels of the language (e.g., colloquial, informal, formal).[8]

---

7 Translation plays an uncertain and paradoxical role in foreign language learning. The goal of foreign language acquisition is comprehension without translation, yet comprehension seems impossible without some word-level translation. The very idea of foreign language translation at all is explored exhaustively by George Steiner (1992). For a more recent exploration of many aspects of language and foreign language translation, see Hof-stadter (1997).

8 See Durrell (1992, p. 3) for a good discussion of register variation in German.

Language students are usually introduced to a foreign language by a text book that justifiably emphasizes the language in its more formal mode. Encountering colloquial, informal, or regional variations can cause confusion. While it is likely that many variants of a foreign language will be encountered at one time or another on the Internet, it is not certain that the proposed system could do any more than flag the more common variants when they are detected. It would be very difficult—and perhaps inadvisable—to attempt to convert variants from one register to another, say from usage acceptable in the popular press to the form that would be used if expressed in a more formal literary medium. For the intermediate-level language student, however, the highlighting of language variants that fall outside the standard, formal register may help to deepen the user's understanding of the foreign language.

4.1.1.2.5 Text-to-Speech

Text-to-speech (TTS) refers to a technology that uses advanced linguistic analysis techniques and synthetic speech generators to convert digitized text to speech. TTS systems can provide surprisingly naturally paced, accurately pronounced, true-to-life inflected, and properly word stressed synthesized speech read-outs of textual material for many languages.[9] The more advanced TTS programs provide output controls for excitement level, overall pitch and gender of the speaker, and volume.

One interesting feature of TTS technology is that it will automatically mirror some of the page layout techniques—such as phase structure spacing—noted earlier. TTS systems artificially produce "natural" speech phrasing, inflection, and word stress based on a linguistic analysis of the textual material. TTS systems grammatically parse textual input and convert it to audio output that is intended to mimic the features common to native-like speech, including phasing.

The availability of TTS technology will ensure that there will always be some minimal audio counterpart for textual selections that lack a readily accessible native-speaker-based audio analogue. The technology can also promote reading comprehension directly by providing the user with an acoustic "parsing" of the text-based material in a way that accords with the linguistics features of the material, confirming or contradicting the reader's tentative interpretive hypotheses. Hearing how a sentence would be read if read aloud is sometimes helpful in disambiguating or untangling the confusing syntax of its printed counterpart.

---

[9]   See Appendix C for a list of some commercially available TTS products.

In addition to the basic manipulation output controls provided with the more advanced TTS programs, the proposed system is to support the full range of listening comprehension controls. The listening comprehension support features of the proposed system are described in the following section.

4.1.2   Listening Comprehension

Natural, fluent, native speech can be thought of as lying at one end of an "articulation continuum"; where "Motherese" or "teacher talk" lies at the other.[10] Good foreign language teachers use a pedagogic form of the language in their teaching. They speak more clearly and slower; they emphasize and exaggerate important words, expressions, or phrases in their target language utterances. They pause longer between sentences and phrases. In doing so, they stretch the fabric of the language, making it more transparent—and comprehensible.

My proposed system is an attempt to mimic this facet of the instructional process but without the teacher. SP capabilities of the system will enable the acoustic enhance necessary to make this possible. They will enable the emulation and replay of a teacher talk-like form of Internet audio material. The system will enable the student to "tune" authentic, native speech to a form that is more understandable; and with comprehension comes acquisition, according to Krashen. Speech processing technology can make this possible.[11]

Support for listening comprehension is to be roughly analogous to that provided for reading support. The user is to be given controls to manipulate the acoustic elements of spoken input to enhance comprehensibility. These controls are to be similar to the techniques envisaged for the reading module.

Listening comprehension support is to be provided by three main presentation controls: transport, phonological, and transcript. The system functionality for listening comprehension

---

[10]   The role of this articulation continuum in foreign language learning is discussed in Delcloque (1995, p. 55). "Motherese" is the special variety of speech that mothers use in talking to their prelinguistic children. Characteristically, it's conversational give-and-take, repetitive drills, and simplified grammar. It actually has little effect on native language acquisition: "Children deserve most of the credit for the language they acquire," according to Steven Pinker (1994, p. 40).

[11] The use of commercial speech processing technology to support foreign language learning is not novel with the proposed system. There are several foreign language learning systems on the commercial market that claim to use speech processing for automatic speech recognition (and subsequent system parsing and error correction), spectrogram comparison of the learner's pronunciation with that of a native speaker, and phonetic transcriptions of foreign language speech. See Appendix C.

support is described in the next subsections and summarized in Table Table 2, "System Functionality for Listening Comprehension Support," on page 35.

4.1.2.1   Transport Controls

The basic audio transport controls of the proposed system are to be similar to those found on standard audio compact disc players used for digitized music replay: start/stop, pause/resume, skip to next selection (in both directions), scan through current selection (in both directions), and repeat capabilities (complete selection, paragraph, sentence, or user-selected subsection). Except for scan and repeat, these transport controls are normally available on sound-equipped desktop computers. In addition, there is to be a paging function, analogous to that provided for text material paging, that allows users to control the rate and "unit" of audio replay.

4.1.2.1.1 Start/Stop

The start/stop control is to be used to initiate and terminate the audio replay of the spoken version of the foreign language input. "Start" initiates audio play from the beginning of the selection. "Stop" functions both to stop the audio output and to "rewind" the material. Rewinding may be the beginning of the complete selection or to the beginning of the current paragraph, sentence, or user-defined subsection (as described in Section 4.1.2.1.5).

4.1.2.1.2 Pause/Resume

The pause control is to be used to momentarily suspend audio output during a listening comprehension session. "Resume" is to be used to continue the audio output from the point of suspension. The "pause/resume" control pair is intended to give the listener an opportunity to temporarily suspend the audio output without, as it were, losing his or her place in the audio stream.

4.1.2.1.3 Skip

The skip control is to be used to quickly move between discrete "sections" of the overall listening comprehension selection. The level of "sectional" granularity of the selection is to be determined and set by the user. A useful level of sectional granularity for a fairly long piece would be the paragraph. For shorter selections, the sentence might be the more appropriate "section" of the material. The user will be able to skip forward and backward through the material, quickly moving to a particular paragraph, perhaps, to resume listening comprehension work that was interrupted earlier.

Table 2. System Functionality for Listening Comprehension Support

Major Function	Sub-Function	Mode	Description
Transport	Start/Stop		Start from the beginning. Stop and return to the beginning.
	Pause/Resume		Pause at current position. Resume from current position.
	Skip	Forward/Reverse	Skip forward/reverse incrementally through selection at paging unit.
	Scan	Forward/Reverse	Scan continuously forward/reverse from current position within selection.
	Repeat [count]	Selection	Repeat selection [count] times.
		Paragraph	Repeat current paragraph [count] times.
		Sentence	Repeat current sentence [count] times.
		A-B	Repeat material between user-specified points A and B [count] times.
	Paging		Analogous to text material paging: "page,"[a] paragraph, sentence, or phrase. Auto-matic, at user-specified interval or on demand.
Phonological	Speech Rate	Global	TSM to vary overall speech rate of audio selection.
	Prosodic:		Variable length pauses between audio elements and/or "enunciation" adjustments.
	Pause Expansion		Exaggeration of inter-paragraph, sentence, phrases.
	Phrase Emphasis		Articulation, pitch, voicing, loudness, and timing adjustments of major phrases.
Transcript	[n] Before		Display corresponding transcript [n] seconds before audio playback.
	Simultaneous		Display corresponding transcript simultaneous with audio playback.
	[n] After		Display corresponding transcript [n] seconds after audio playback (shadowing) .

a. Audio "page" is defined by corresponding text page (screen), if available.

35

4.1.2.1.4 Scan

The scan control is to be the analogue of fast-forward and fast-reverse controls on audio tape players. Unlike the skip control, it will allow the user to move in a continuous as opposed to a discrete fashion within the material. It is likely to be used to quickly "back up" for immediate replay. Typically, this rewind distance would be smaller than the value in effect for the skip control.

4.1.2.1.5 Repeat

The repeat control is to be used to automatically "rewind" and replay the audio selection either when the end of the selection has been reached or when the user invokes the stop function. Repeat modes are to include the complete selection (from the beginning); a standard subsection (paragraph or sentence); or an arbitrary region previously set by the user by specifying the start (A) and stop (B) points. This "A-B" repeat function could be used to specify phrases within a sentence, multiple sentences, or multiple paragraphs for automatic repeat.

It is to be possible to specify the number of times the material is to be repeated before the system automatically advances to the next section of the selection. In this way, a user would be able to listen to three replays (let's say) of each sentence or paragraph of the source material before advancing automatically to the next sentence or paragraph.

4.1.2.1.6 Paging

An audio playback paging function, analogous to that provided for text material paging, allows users to control the rate and audio "unit" for replay. The rate of replay is determined by the user-specified interval between successive "units" of audio material. Alternatively, paging can be on demand. The audio unit is either a "page" (defined by the corresponding text page or screen, if available), a paragraph, a sentence, or a phrase.

4.1.2.2 Phonological Controls

Phonemes are the basic units of speech. They correspond roughly to the letters of the alphabet. In combination they compose spoken words. Phonology is the study of these fundamental speech elements, how they combine, and how they are used to produce natural speech. The ability to detect the phonemes of a language and to discriminate between them is a critical first step to speech understanding.[12] The phonological controls of the proposed system are aimed at producing an environment that promotes those language comprehension processes that depend (at least in part) on primarily phonological factors. While the transport controls

provide the user with control over *what* is played, the phonological controls give the student control over *how* the material is acoustically "formatted." They give the user useful control over the phonological environment. Two sets of such controls are envisaged: speech rate and prosodic.

4.1.2.2.1 Speech Rate Control

The proposed system is to provide a global speech rate control using time-scale modification (TSM) speech processing technology. TSM technology offers the capability to replay audio material at rates slower or faster than the original recording while preserving, within limits, the "natural" sound quality of the original voice. TSM is commonly used in speech research and for fitting audio segments to required intervals, for example, in radio or television commercials.[13]

TSM is to be used for variable-rate playback in the proposed system, allowing the student to slow or increase the rate of playback without perturbing the fidelity of the original audio. Typically, the user will "expand" the audio signal, effectively slowing the speech rate and giving the student more time to comprehend what is being said. Speech rate control is to be exercised globally, that is, to the entire selection.

4.1.2.2.2 Prosodic Control

The second set of phonological controls is intended to use "prosodic" techniques to clarify the syntactic structure of the spoken material.[14] These techniques are to include pause expansion and phrase emphasis.

Pause Expansion

Pause expansion is another technique for controlling speech rate. But unlike TSM, which is to be used to slow or speed up the overall speech rate, pause expansion is to be used lengthen or exaggerate the natural pauses that occur between paragraphs, sentences, and individual phrases or clauses within a sentence. The intent is twofold. First, it is to give the lan-

---

[12] Recognition of phonemes qua phonemes is only part of the language acquisition story. According to the neuro-philosopher Daniel Dennett, "[t]he segmentation of speech sounds [i.e., of the phonemes and words composed of phonemes] is a process that imposes boundaries based on the grammatical structure of the language, not on the physical structure of the acoustic wave" (1995, p. 51).

[13] See Appendix C for a list of some commercially available time-scale modification products.

[14] "Prosodic" is being used here in a very general sense to refer to all patterns of articulation, pitch, voicing, loudness, and timing found in speech. In this sense, prosodic elements play an important role in conveying the meaning of an utterance.

guage learner more time between phrases or sentences to try to understand what was just said before being quickly forced to deal with the next phrase or sentence. Second, the purpose is to help clarify the structure of the selection, particularly the phrase or clausal structure of each sentence.

The system will support pause expansion at the paragraph, sentence, and phrase level. with different pause intervals specifiable for use between paragraphs, sentences, and phrases. Typically a longer pause will be used between paragraphs than between sentences within a paragraph. The pauses used to clarify the phrase or clausal structure of a sentence will be only slightly exaggerated and shorter than those to be used between sentences.

Phrase Emphasis

In addition to pause expansion used at the phrase level, the proposed system is to support prosodic enunciation to emphasize important phrases or other elements of the spoken material. Typically the student will want to have subject phrases emphasized in order to help in the quick identification of the nominal topic of each sentence. Emphasis can be achieved via simple amplification. The intent is to approximate the "teacher talk" used by skilled and experienced foreign language teachers to convey meaning by emphasizing the important elements of the sentence.

4.1.2.3   Transcript Controls

In many instances, a written transcript will accompany the foreign language audio material available on the Web. If such a transcript is available, it can be downloaded and used to complement the listening comprehension task by allowing the user (optionally) to "read along" as he or she is listening to the audio. The proposed system is to provide several transcript control features to be used in conjunction with the audio transport and phonological controls described above.

Similar to the paging controls for reading comprehension, the transcript display size is to be set by the user at the page (window), paragraph, or sentence level. The transcript display will track the corresponding segments (page, paragraph, sentence) of the audio playback. Moreover, the interval between the display of the transcript and the playback of the corresponding audio segments is to be adjustable and will allow for the display of the transcript before, simultaneous with, or after the playback of the corresponding audio signal. The interval adjustment can be in terms of a fixed amount of time after the beginning of the audio stream or immediately after natural break points such as the end of a paragraph or sentence.

The "shadowing" effect obtained by the display of the corresponding transcript a few seconds after the audio segment has been played permits the user to focus on the listening comprehensive task but with the benefit of immediate confirmation of what was heard by consulting the transcript.

The transcript display is also to provide a visual (typographic) analogue of any phonological modifications introduced into the audio stream. So, for example, if the user had requested an enunciatively emphasized form of all subject phrases, the corresponding textual form would display subject phrases appropriately emphasized (e.g., using bold-face or italics).

For audio material that is not accompanied by a word-level transcript, ever improving speech recognition technology could be employed in the future to automatically generate the needed transcript.

4.2   Classroom and Research Use

The proposed system was originally conceived as a learning tool for independent use by individual language learners outside the classroom. It is not hard to see, however, how the system could be employed for both pedagogic and research purposes within a more traditional language learning environment.

4.2.1   Pedagogy

While the design of the proposed system is mainly focused on the individual language learner, the system can exploit its two important core technologies—NLP and SP—for pedagogic purposes. The ability to linguistically parse authentic or textbook material and project the results of that parse to an entire class to illustrate some grammatical point would be a powerful teaching aid. Probably the best use of the system would be in its use to teach grammar, providing the teacher with a tool to illustrate the variety of grammatical constructions from both authentic and textbook material.

More in line with the intelligent tutoring approach to language learning, the system could be augmented with a tutorial that incorporates any sound theory of the L2 acquisition process and which can automatically lead the student through the natural stages of listening and reading comprehension. The tutorial—like any talented foreign language teacher—would modify the presentation based on knowledge of the individual learner's learning style, proficiency level, type of material being presented, and computer-diagnosed student weaknesses.

39

4.2.2 Research

If used in the classroom, the system could also serve as a powerful research tool, capturing learner usage data in a log file from which second language acquisition researchers could investigate directly some of the metacognitive aspects of the L2 listening and reading comprehension process. For listening comprehension research, the system could keep track of the number, size, and type of audio replays as well as the use of the particular audio modulation devices (speech rate changes, lengthened pauses between major phrase, sentences, etc.) invoked. Analogous tracking of the linguistic cues highlighted by the student would also be clearly valuable for reading comprehension research.[15]

---

[15] This idea was originally advanced by Nina Garrett (1990).

Chapter 5. System Architecture

The purpose of this chapter is to sketch a system architecture for the proposed system. The major components of the suggested architecture are available today. Clearly, more work needs to be done to verify the technical feasibility of the approach I am presenting. I am not aware of anyone who has attempted to cobble together SP, NLP, Web browsing, and text formatting tools in the way I suggest. The next step is to built a proof-of-concept prototype that demonstrates the feasibility of the proposed system.

5.1   General System Configuration

The proposed system is envisaged to operate as a Web-browser "plug-in" application. This plug-in application is to preprocess Web-based foreign language text and audio source material for subsequent display in accordance with the preferences and language learning needs of the user. For text material, the system would be configured to lie directly between the user's Web browser and the user. The browser would be used both to search the Web for interesting, target language material and to subsequently retrieve and parse hypertext markup language (HTML) based text pages.[2] The resulting text would then be automatically piped to the proposed system for linguistic analysis, reformatting, and display to the user.

For audio material, the system would reside between an audio "viewer" that is itself a browser plug-in[3] and the user. The audio play plug-in would be responsible for converting the audio material from the usually proprietary and compressed form in which it is transmitted over the Internet to a form that can be sent through a digital signal processor for playback to the user. Instead of being fed to the audio card and speakers on the user's computer, however, the audio signal would be fed into the proposed system for further processing before playback. The general system configuration for both text and audio input is illustrated in Figure 4.

---

[2]   The system does not require HTML-based input. Documents encoded in the portable document format (PDF) are also widely available on the Web. They would be equally amenable to the linguistic pre-processing and reformatting techniques of the proposed system.

[3]   RealPlayer Plus 5.0 from RealNetworks (formerly Progressive Networks) is a viewer application designed for display of Web-based audio, video, and animation material. Available at www.real.com.

HTML-based
Text
Pages

Plain Text

Text

Web

Web-Browser

Foreign Language
Listening and
Reading
Comprehension
System

User

Audio Player
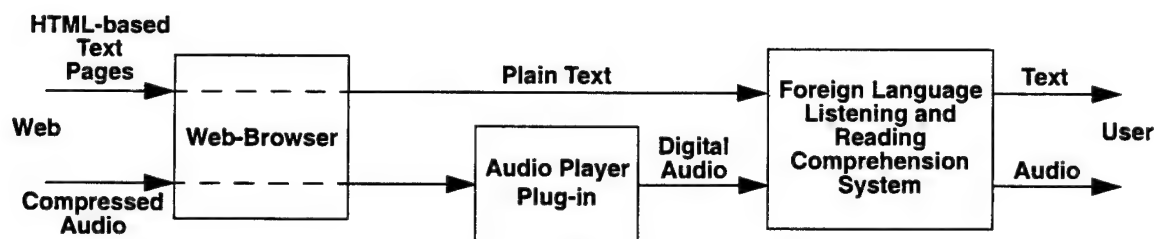Plug-in

Digital
Audio

Compressed
Audio

Audio

Figure 4. System Configuration

It should be noted that the configuration suggested in Figure 4 is highly simplified. Common Web browsers do not simply convert HTML-tagged text to "plain text." Instead, they display or "render" the textual material is accordance with the specifications of the HTML tags (e.g., "<B>text</B>" for bold-face, "<Center>text</Center>" for centered text), the text display capabilities of the platform on which the browser is running, and the built-in rules of the browser for interpreting HTML-tags. The HTML-based text is displayed on the user's workstation screen with as much format desired by the page author as possible within the limitations imposed by the user's display screen (e.g., current window size, graphics capabilities) and the user's browser. As much of this formatting information as possible should be preserved as it passed into the proposed foreign language learning system. Hyperlinks especially ought to be preserved as Web pages are passed from the Web browser to the proposed system.

Accordingly, it may be simpler and more desirable to have the Web browser pass the HTML-based pages to the proposed system directly without interpretation of the HTML tags. The foreign language system would then be responsible for preserving the HTML tags and using them—consistent with its own reformatting of the input text stream—in its final display to the user.

5.2 General Architecture

The proposed system is to consist of two basic components: a text and audio material preprocessor and a display processor. The preprocessor is responsible for the linguistic encoding of the audio and text source material prior to display. The display processor displays the linguistically encoded material in accordance with the display specifications defined by the user. It also synchronizes text display with audio replay.

42

5.2.1   Preprocessor

The preprocessor will have to accommodate three different input possibilities: (1) the availability of both text and audio material, (2) text only, and (3) audio only.

The preprocessor configuration that is necessary when both text and audio material are available is illustrated in Figure 5.
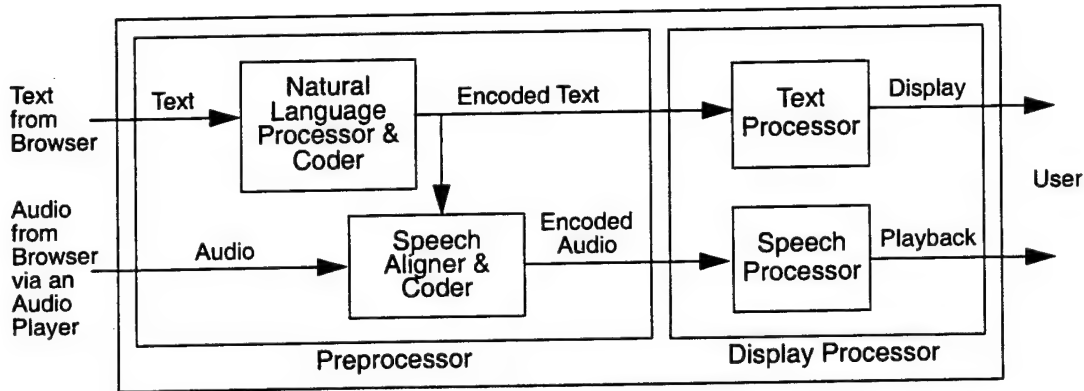

**Figure 5. System Architecture for Both Text and Audio**

Under this scenario, both audio material and an accurate word-level transcript of the spoken material are available from the selected Web site. The transcript is parsed by a source language NLP application to determine the linguistic structure of the text. The results of this linguistic analysis are then encoded to enable a subsequent formatted text display (by a display processor) in accordance with the user's display specification. Encoding is likely to take the form of embedded markers that indicate for each morpheme, word, clause, phrase, and sentence the linguistic role that linguistic element plays within the material.[4] Each marker would indicate the linguistic role the element(s) plays within the material.

The encoded text may then be used to support the preprocessing of the audio signal. The availability of an accurate word-level transcript simplifies the speech processing task, enabling a speech aligner to more easily determine word and sentence boundaries in the audio signal. The audio material is further encoded using the results of the linguistic analysis conducted on the textual form of the material. This will guarantee that any replay of the audio version of the

---

[4]   For performance purposes, the extent to which the material is linguistically parsed and then encoded may be a function of the user's predefined display specification. Although morphological analysis is necessary for higher-level syntactic analysis, full morphological detail need not be retained as part of the encoded text for text display if the user is not interested.

43

material that emphasizes some one particular syntactic feature will track the corresponding display of the text version of the material with the same feature highlighted.

In situations in which only text is available, the proposed system will have to rely on speech synthesis technology if it is to generate an audio version of the text. The architecture needed to accommodate this situation is illustrated in Figure 6.
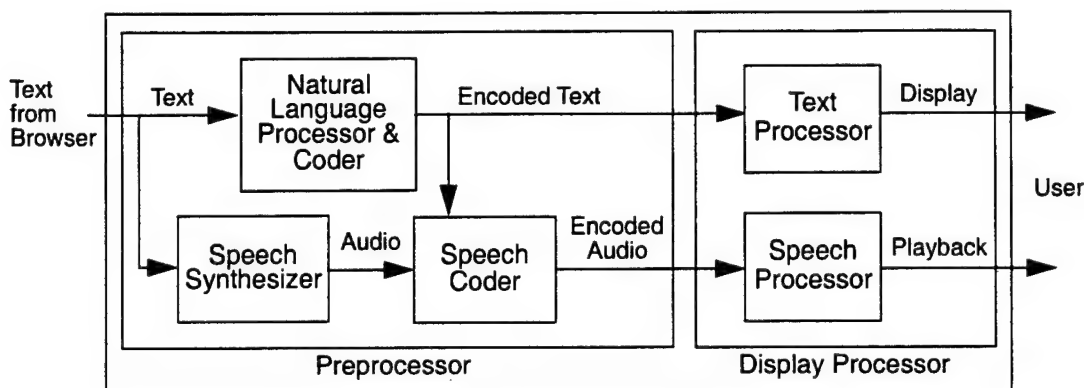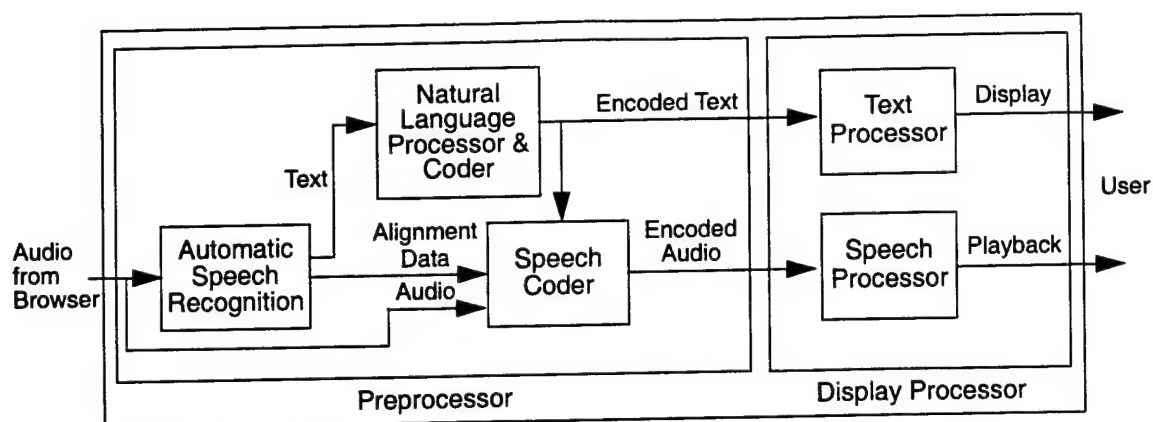

Figure 6. Text-Only Architecture

There are two important differences between the text-only architecture and the system architecture pictured in Figure 5. First, in the text-only architecture an audio signal has be generated by a speech synthesizer. This complicates the basic architecture but provides in return greater flexibility. Audio playback can be provided for the Web sites that do not provide audio. But as a consequence of this complication, subsequent audio preprocessing is simplified. This is the second important difference. There is no longer a need for a speech alignment unit since the audio has been generated directly from the text input. It already knows where the word breaks lie.

For situations in which only an audio stream is available, the proposed system will have to rely on automatic speech recognition to generate both a text version of the material as well as the encoded audio signal for subsequent display.[5] This audio-only architecture is illustrated in Figure 7.

---

[5]   For the kinds of speech processing requirements discussed in this paper, a number of commercial products are available from different companies, including Entropic Research Laboratory, Inc., SRI International's Speech Technology & Research Laboratory, and InfoSignal, Inc.

Figure 7. Audio-Only Architecture

There is considerably more complexity in Figure 7 than in either of the two previous figures. Moreover, this architecture is only one possibility resulting from somewhat arbitrarily allocating necessary preprocessing functionality to individual subcomponents. The Alignment Data arrow, for instance, is intended to show how information obtained from the ASR unit could be used to simplify the subsequent speech coding task. Alternatively, an integrated Speech Aligner and Coder component (not shown) could have been proposed to preprocess the raw audio input independent of any prior ASR processing. Similarly, all linguistic analysis functions have been relegated to the NLP and Coder component even though a considerable amount of linguistic analysis is involved in the ASR-production of the text input to the NLP and Coder units. The linguistic analysis assumptions upon which the ASR result is based could be used to make the NLP component more robust.

5.2.2 Display Processor

The role of the display processor is to format and display the text and audio stream in accordance with the display specifications provided by the user. In the presence of both text and audio, it must also synchronize the text display with the audio signal, ensuring that the text stays current with the audio version of the material.

The basic form of a display specification is a set of ordered pairs, $\{<g_1, f_1>, <g_2, f_2>, ... , <g_n, f_n>\}$, where the g_i's denote a grammatical construct of interest to the language learner and the f_i's denote a formatting technique. The user indicates an interest in passive voice constructions, and also indicates that all such instances occurring in the input material should be displayed in bold-face type. The display processor is responsible for the capture and mainte-nance of the display specification and for its application to the display output. Conceptually,

45

the display processor will examine the grammatical codes of the encoded input for possible matches with the g_i's of the display specification. Finding a match, the display processor will format that element of the output in accordance with the corresponding format specification, the f_i's.

But how will the user build the display specification? It should depend on what the user is trying to do. It will also depend on the grammatical constructs used in the target language; these differ, of course, from language to language.[6] If the user is interested in seeing all instances of a certain construction throughout the material, a static, non-contextual menu-based approach is certainly acceptable. A cascading pull-down menu illustrated in Figure 8 would do.
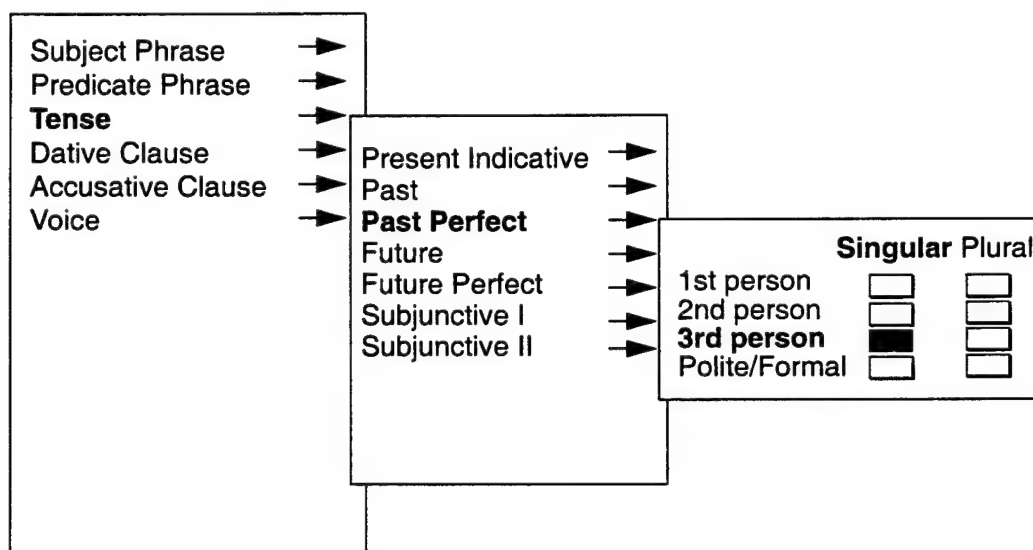
Figure 8. Cascading Pull-Down Menu to Select Language Feature

The user would simply pick one or more items from a standard list of grammatical features for the given target language, along with a format specification for each selected item. The example in Figure 8 shows the user selecting "tense {past perfect {3rd person singular}}."

For the more advanced language student, however, I would recommend a more dynamic, context-based approach that is tailored to the features available in a specific selection of the input material. Take the German sentence *Die Strasse, in der er wohnt, ist sehr elegant*. The segment *in der er wohnt* contains (1) a relative clause (*in der*), (2) a relative pronoun (*der*), (3)

---

6   Note that this means that the user interface will have to be customized to reflect both the user's native language and the target language. The German language examples used in the paper could be used with only minor changes in a version of the system for native French speakers; only the interface would have to be translated into French.

46

a dative-case prepositional phrase (*in der*), (4) a "two-way" preposition[7] (*in*), (5) a possibly confusing example of dependent clause word order, or (6) just the present third-person singular form of a typical regular (often called "weak") German verb (*wohnen*). Adjective endings, however, play no role. If a user highlighted the phrase, the system should limit the menu options to those listed above, ignoring adjective endings, subordinating conjunctions, the subjunctive, the passive, and many others.

By the way, this example illustrates one drawback of some language learning software currently available on the commercial market. One popular program gives the user a running account of the grammatical form exhibited by each word or "segment" of a given canned selection or "title." This is actually very nice and occasionally useful. It is limited, however, to the grammatical forms that the title author chose to use to describe the grammatical function the word or phase plays within the encompassing sentence. The simple, two-word phrase *in der* would likely be correctly identified as a relative clause, but the fact that it also illustrates the use of a German two-way preposition, and, in particular, the dative form of that preposition, will probably be omitted. This is simply because the linguistic analysis that goes into the authoring of such products is done largely by hand. A human must painstakingly mark-up each word and segment in the selection, encoding each word and segment's grammatical role in the material.[8] The result is a product that is limited in its flexibility for displaying grammatical features of interest to any student, at any level. The products themselves are also limited in the breath and range of authentic material they can and do make available at an affordable price.

---

[7] Meaning it can take either the dative or accusative case, depending on how it is used in the sentence.

[8] This is not entirely a manual process. There are tools available to facilitate language structure mark-up for multimedia authors. See, for example, Lyman-Hager (1995).

Chapter 6. Discussion and Summary

Until now I've intentionally avoided some of the criticisms that might be leveled against the proposed system. Now is the time, first to address the skeptics and then to draw the paper to a close by summarizing the key features of the system.

6.1   Issues

Is there any evidence that the proposed system will work? Or should the system itself perhaps be construed as the mechanism by which to test the (implicit) hypothesis that the artificial emphasis of linguistic cues can promote foreign language comprehension and learning? This hypothesis, after all, was one of the four key *assumptions* in Chapter 3 on page 17. I also alluded to the research possibilities of the system at the end of Chapter 4, page 40.

There are actually two different forms of this fundamental assumption. On the one hand, the system assumes that typographical techniques can be used to promote comprehension of foreign language *text*. On the other hand, I'm assuming that *acoustic* manipulation can promote comprehension of foreign language *speech*. The two techniques are not the same nor are they necessarily directed at engaging the same cognitive mechanisms. It's possible that typographic enhancement of text may promote comprehension while acoustic manipulation of speech has only negligible effect, if any, on language understanding. Or both techniques may fail.

6.1.1   Typographical Techniques

I am not aware of any formal studies that attempt to quantify the efficacy of typography on written language learning. Grade school reading materials are printed in large, unadorned, and non-ornamented type fonts with generous word and line spacing. Along with simplified grammar, short sentences, and the use of common, simple, concrete, and everyday words, the typography of these materials must surely play an important role in learning to read. The publishers of pre- and grade-school reading primers are intuitively leveraging a widespread behavioral phenomenon known as *supernormal stimulus*. The Harvard biologist E. O. Wilson

describes this as the innate "preference during communication for signals that exaggerate the norms..." (1998, p. 231). The beginning reader is innately attracted to anything that differs from the norm or in any way stands out.

Unlike the acquisition of spoken language, the beginning reader has to *learn* to distinguish not only individual words but clauses, sentences, and paragraph boundaries as well. These boundaries are an important part of the linguistic scaffolding that underpins our written language. Designers of reading primers are aware of the importance of this scaffolding and use it effectively in conjunction with a broad array of other typographical techniques to promote the reading learning process.

The adult learner has long since automatized the ability to recognize word, sentence, and paragraph boundaries. If the foreign language shares a basic typography (alphabet, punctuation marks, left-to-right, top-to-bottom line continuation) with the learner's native language, some of the basic page-layout techniques (larger and simpler font, increased word and line spacing) are probably of little value. Grammatical differences between the adult learner's native language and the foreign, target language, however, take on a new significance. Word order matters in English; it's less important in certain case-based languages. The adult, English-speaking second language learner now has to learn to pay more attention to case markers and to give up undue reliance on word order. Simple typographic devices can play the same role in teaching beginning adult foreign language readers to pay attention to case markers as word and line spacing plays in teaching the child reader to pay attention to words and sentence-terminating punctuation. Each set of techniques serve to remind the beginning reader that certain syntactic features of the text are critical to comprehension. They must be *noticed* and typography is one good way to make sure they are noticed.

6.1.2 Speech Processing Technology

What about speech? Is there any evidence that the use of speech processing technology can improve listening comprehension? Yes. There have been several recent studies that show that (native) language comprehension in language learning impaired (LLI) children can be improved with acoustically modified speech (Tallal et al. 1996; Merzenich et al. 1996).

Paula Tallal and her colleagues have suggested that certain reading disabilities such as dyslexia may be the result of earlier failures in natural language development when exposed to native speech. They hypothesized that these developmental problems may be caused by basic phonological processing deficiencies, specifically a "basic deficit in [the ability to] rapidly [process] changing sensory inputs." It seems these LLI children "commonly cannot identify

50

fast elements embedded in ongoing speech that have durations in the range of a few tens of milliseconds, a critical time frame over which many phonetic contrasts are signaled." The synthetic extension in time over which these elements are signaled, it was conjectured, might allow these same children to make the necessary discriminations. Tallal and her colleagues then went on to predict:

> If the critical acoustic cues within the context of fluent, ongoing speech could be altered to be emphasized and extended in time, then the phonological discrimination and the on-line language comprehension abilities of LLI children should significantly improve.

This prediction was tested by comparing speech discrimination and language comprehension abilities of LLI children both before and after training with synthetically modified input speech. The input speech on which the children were trained was synthetically modified in two ways. First, the duration of the speech signal was prolonged by 50 percent, but with adjustments to maintain the quality of natural speech. Second, speech elements involving very rapid frequency changes (3 to 30Hz) as in the syllables [ba] and [da] were played by as much as 20dB louder. The result was recorded speech that "had a staccato quality in which the fast (primarily consonant) elements were exaggerated relative to the more slowly modulated elements (primarily vowels) in the ongoing speech stream." The researchers reasoned that the amplification of the fast elements in conjunction with overall temporal extension would render these elements more salient and result in a generally "more sharply modulated form of speech." This more sharply modulated form of speech should, in turn, promote the complex signal learning that is listening comprehension.

The studies that tested this prediction were very positive. After only one month of daily training with acoustically modified speech, LLI children showed significant improvement in speech discrimination, language processing, and grammatical comprehension. LLI children who were between one and three years behind their chronological age in speech and language development showed significant improvement, with each of the LLI children in one study "approaching or exceeding normal limits for their age in speech discrimination and language comprehension."

I have been suggesting that similar techniques—time-scale modification, phrase emphasis (via amplification)—can profitably be used to promote foreign language listening comprehension. While the adolescent or adult foreign language learner is certainly not in the ordinary sense language learning impaired, he or she confronts much the same problem: how to effectively process "rapidly changing sensory inputs." To the foreign language learner, the

51

phonetical streams of an unfamiliar foreign language surely present the same kinds of sensory input processing problems that confound the LLI child. The proposed system was conceived specifically to enable the language learner to control this sensory input in a constructive and comprehension-promoting way.

6.1.3   Comprehension Side Focus

The proposed system is limited to the comprehension side of the linguistic equation: reading and listening. As such, it addresses only two of the four classically defined skills of language ability, omitting writing and speaking. Thus it intentionally omits language production. This omission is intentional for two reasons. First, according to Krashen's "Comprehensibility Input Hypothesis," comprehension is logically prior or, in some sense, more fundamental to language acquisition than is language production. Language acquisition begins with input— somewhere between 18 and 24 months-worth in children—before giving rise to (grammatical) output or production. Comprehensible input, as noted in Section 3.2.1 on page 17, is a necessary condition for language acquisition. Its logical primacy is recognized. Accordingly, the proposed system also stresses this aspect of the language acquisition process.

Second, the use of computer technology—e.g., artificial intelligence, multimedia, the Web—to effectively promote foreign language *production* skills is just too difficult. It is hard enough to produce a good syntactic parse of grammatically correct foreign language sentences produced by native speakers. To parse and be able to respond intelligently—for that's what such a system would have to do—to a language learner's foreign language input, either written or spoken, requires the ability to both detect and then correct an immense range of possible errors, errors in both syntax and pronunciation. Accordingly, commercial language tutoring systems that focus on production skills are usually very constraining in what they allow in the way of user input. Users "fill-in-the-blanks" and the system checks for syntactic correctness. This is a far cry from a genuine ability to produce meaningful and understanding utterances in a foreign language, as any student or teacher of a foreign language will attest.

While a reasonable case can be made for limiting the system to comprehension skills, there is nevertheless a certain awkwardness to the system's listening-reading pairing in contrast to possible listening-speaking or reading-writing parings. Our instinctual or innate preliterate language ability is verbal and auditory. Normal humans reared in ordinary circumstances can't help but acquire language, and at a very early age. Indeed, children "aren't really learning [their mother language] at all, any more than birds learn their feathers" (Dennett 1995, p. 388).[2] On the other hand, reading and writing abilities are not instinctual but *learned* cultural skills. They

reflect higher cognitive abilities. While Krashen's theory on language acquisition may accord with the instinctual acquisition of verbal linguistic ability, it is not at all clear that it applies equally well to learned, non-verbal linguistic ability (i.e., literacy). Lots of comprehensible *verbal* input may be a necessary condition for natural (verbal) language acquisition; it doesn't follow, however, that lots of written input is a necessary or even catalyzing condition for the learning of a written language. Also contrary to the spirit of Krashen, the written comprehension sub-system exclusively emphasizes the strictly linguistic features of the target language. This emphasis on grammar—what Krashen disparages as "the structure of the day"—is often blamed for the lack of enthusiasm and hence success in foreign language learning. One has to learn to read for "comprehension," the argument goes, by going beyond the grammar that often stands in the way of real understanding. Focusing too much on surface forms obscures the underlying meaning and actually hinders comprehension.

Is this attack on the theoretical underpinnings of the system a fatal objection? No. If Krashen's "Comprehensible Input Hypothesis" applies only to the acquisition of spoken languages, then a system that attempts to enhance comprehension—and hence learning—of a written foreign language based entirely on the provision of "lots of comprehensible input" is indefensible on SLA theoretical grounds. But the amount of textual input that can be processed and provided by the proposed system is only one feature—and perhaps the least important feature—of the system. Its distinguishing and important characteristic is that it can promote the detection of purely linguistic cues to facilitate learning of the written and purely conventional form of the target language. Written language is a cultural artifice. It has to be learned; it is not an instinctual process. Rather than being out of place in promoting written language learning, techniques that can render more clearly the forms and purposes of the purely surface-level elements of the language have to have a positive effect on the learning process. Learning to read, whether one's native language or a foreign language, is just learning to detect the many surface-level cues that collectively enable meaning to be conveyed from a writer to his or her reader.

I see no inherent conflict or inconsistency, then, in a Krashen-inspired system that promotes both listening and reading comprehension. The techniques I've described that could be used to promote comprehension in both verbal and literal comprehension skills are similar: acoustic techniques are proposed for the auditory sphere, and typographic devices are suggested for the written arena. The proposed typographic techniques are potentially more powerful; I see no easy and non-distracting way to flag idioms or proper nouns, for example, using acous-

---

2   See Pinker (1994) for an engaging account of (Chomsky's theory of) language acquisition as an innate facility of the human mind.

tic devices. But the two sets of techniques really are focusing on different problems in foreign language comprehension and hence foreign language learning. The acoustic techniques are basically there to allow the learner to slow things down and to impart phrasing that may make things easier to understand. Acoustic emphasis of case markers, on the other hand, to clarify the grammatical function of certain words or clauses just doesn't make any sense. Highlighting of case markers in textual material, however, can be very helpful, especially if their use differs considerably from that of the learner's native (written) language. At a certain point, the typographical devices may begin to interfere with the learner's grasp of the material. At that point, the learner can and should turn them off for he or she has reached his or her goal.

6.1.4   Contrast with Commercially Available Products and Tools

I am not aware of any commercial foreign language learning products that attempt to provide the full range of features being proposed in this paper. There are foreign language translation products widely available on the Internet. They enable Web users, for instance, to get a machine translation of foreign language Web pages. One popular such service is provided by Systran Software, Inc., and is linked with the AltaVista commercial Internet search engine. Web pages for which a translation can be obtained are indicated with a "Translate" flag on the Uniform Resource Locator (URL) line of the list of results returned by AltaVista. For example, the following URL appeared below one result returned by AltaVista on a Web search for "Goethe":

http://www.goethebuch.de/ - size 3K - 22-Jan-98 - German - Translate

It indicates that the Web page is in German and that a translation service is available for the document by simply clicking on the hyper-linked word "Translate." Clicking on the highlighted link would bring up a dialog-box that requests the user to specify the source-target language specification for the translation. Since the example document is in German, if I wanted an English translation, I would select the option "Translate from German to English." If the page being requested was large, only the first paragraph or so is actually translated for you. And all requests to retrieve subsequent or linked pages from the translated page requires the Web user to go through the "select a translation specification" dialog-box. Tedious, at best.

But we are not interested in translation of foreign language Web sites. The proposed system is intended to use the results of the intermediate step in a Web page translation service—namely the initial linguistic parse—to re-format the page in a manner that enhances comprehension. The proposed system does not need to translate the foreign language Web page; it needs only to correctly identify the syntactic features of the material in order to be able to

emphasize features that the user wishes to have emphasized. The existence of Web page translation services on the Internet thus demonstrates the technical feasibility of the parsing and re-formatting of Web page material for the user in a manner similar to that being proposed with the current system.

While Web page foreign language translation services reflect one important feature of the proposed system (pre-processing of Web-based, authentic foreign language material "on the fly"), they do not really get at the feature that I consider the essence of proposed system: the emphasis of the syntactic cues that facilitate and promote language comprehension. Again, to the best of my knowledge, there are no commercial products that try to exploit the potential of NLP and multimedia for foreign language learning in the way I've outlined in this paper. The various CD-ROM titles from Transparent Language's *LanguageNow!*, however, are moving in the right direction.

The Transparent Language CD-ROM titles provide the foreign language learner with both synthetic and authentic foreign language text. The authentic material is taken from literature, travel brochures, magazines, and so forth. The text is displayed in one window, scrolling to follow a native speaker's audio recitation of the material. A series of "full-motion video-like" frames that portray and follow the action of the script is presented in a smaller video window. This video window can be toggled open or closed. A window containing a English translation of each "segment," generally a complete sentence, is provided immediately below the window containing the foreign language text. This translation window can also be toggled open or closed. There is yet another window that provides a grammatical "analysis" of the words or phrases that the user selects for "analysis." This grammatical analysis typically displays the part of speech of the selected word or, if it's a verb, the particular tense along with its person and number. Like most of the other features of the program, this grammatical analysis window can also be toggled on or off.

There are a couple of interesting controls that *LanguageNow!* offers to support reading and listening comprehension. The user can select either a word, a segment, or the complete selection (which may consist of 100 or so continuous segments) for display and playback. The audio playback rate can be adjusted without any noticeable loss in authentic-sounding speech. There is even a facility that allows the user to record his or her pronunciation of a word or segment of the selection and then compare the student's and native speaker's waveforms on several pronunciation aspects. These include articulation, pitch, voicing, loudness, and timing.

In general, the *LanguageNow!* program is an impressive multimedia-based foreign language learning product. Some of its features are similar to those proposed in this paper; in those

instances of very similar functionality, however, the system being proposed here is far more powerful, flexible, and general. There is no way to automatically repeat any given selection of the material in the *LanguageNow!* product, for example. An individual word (or segment) can be repeated but only by first selecting the word or segment and then repeatedly clicking on the "speak" button. In my proposed system the user would select any arbitrary continuous section of text, select a repeat factor, and then have the selection repeated automatically for the specified factor. Or the user could specify that the system automatically repeat each sentence or paragraph n times, again without further user intervention.

Nor does *LanguageNow!* provide those features that I think are critical for successful (computer-assisted) foreign language learning. Although most of the material is authentic, the system cannot provide the copious amount of material that Krashen and others feel is necessary for effective learning. The number of separate titles is around 25 per language.

The grammatical analysis is canned. The material is grammatically "marked-up" prior to publication. It reflects the perspective of the "teacher" and not the learner. As I noted in Section 5.2.2 on page 45, any given sentence (or phase) can exhibit several different grammatical aspects at the same time. Any pre-defined or "canned" analysis necessarily limits the user to that grammatical aspect that was more or less chosen arbitrarily by the person who prepared the original mark-up. That aspect may or may not be of interest—and hence effective—to the learner.

In contrast to existing commercial products, the system being proposed here returns control to where it belongs: to the user. The user is to select those grammatical features of the target language in which he or she is interested and then let the computer detect and highlight all tokens of that type found in the text. The process is intended to enable the student to focus on just those syntactically aspects of the language that are causing problems or are of most interest. And, because those problems will vary from student to student or even with respect to the same student at different stages in his or her individual learning process, the system must be flexible enough to allow a change in focus at any time. The system being proposed here provides this ability to re-focus the learner's attention to those $n+1$ linguistic constructs that are the hallmark of successful foreign language acquisition.

6.2   Summary

There are many multimedia-based foreign language learning products on the market. In my opinion, none is as firmly grounded in SLA theory as is the system being proposed here. While they all, for the most part, use powerful and sophisticated multimedia technology, they

do not apply that technology to deliver one of the most important aspects of second language learning: copious quantities of comprehensible authentic foreign language material. The system being proposed is designed to take advantage of the enormous quantity of such material now readily available on the Web. It is designed to enable the formatting of that material in a way that can promote both listening and reading comprehension. It returns control of the language learning process to the student in the belief that the (intermediate level) student knows what aspects of the target language are causing the most problems. The system leverages three converging technologies: multimedia, artificial intelligence, and the Internet. These technologies are employed in the service of aspects of contemporary SLA that seem to have the most efficacy in successful foreign language learning.

The proposed system can enhance the foreign language capabilities of DoD. A proof-of-concept prototype could demonstrate both the technical feasibility and genuine power of the basic approach outlined in this paper. The techniques described in this paper for the use of multimedia, artificial intelligence, and Internet technologies for language learning will soon become an important—and commonplace—part of DoD's foreign language instructional enterprise.

Appendix A. Examples

This appendix presents a series of figures that illustrate several of the intended features of the proposed system. Most of the examples use color and bold-face type to bring grammatical features of the material to the individual reader's attention. The particular features of material being reformatted and the particular reformatting device used are noted in a small legend box in the lower right-hand corner of each figure. The examples are limited to the reading comprehensive aspect of the proposed system.

Figure A-1 represents the unmodified source input material that is variously reformatted in subsequent figures in the appendix. The source input is a news story from the German newspaper *Die Welt* (www.welt.de). The story was downloaded from the Internet via the Netscape browser. (The news story is announcing the final agreement among the 16 German Länder (States) of the German spelling system reform, the so-called *Rechtschreibreform*.)

Figure A-2 illustrates the use of color to identify each subject phrase of the source material. Each subject phrase is displayed in blue.

Figure A-3 illustrates the use of color to identify the verb forms used in the article. Each verb form is displayed in green.

Figure A-4 illustrates how the proposed system could be used to help a student understand the differences between the German passive voice and the German future tense. Both the passive and the future tense in German use the auxiliary verb *werden* (to become). The future tense requires the infinitive form of the main verb. In contrast, the passive takes the past participle. The single instance of the future tense form is highlighted in red. All instances of the passive voice are highlighted in blue. The fact that there is only one occurrence of a future tense form in this article underscores the fact that the future tense is far less common in German than in English.

Figure A-5 illustrates how German prepositions can govern different cases, that is, require the use of different case markers on articles and adjectives. Prepositional phrases that govern the dative (indirect object) case are shown in blue. The preposition itself and all corre-

sponding case markers are displayed in bold-face. Prepositional phrases that govern the accusative (direct object) case are shown in red. Again, the preposition itself and all corresponding case markers are displayed in bold-face. German "two-way prepositions," (i.e., those that can take either the dative or accusative case, depending on the role they play in the sentence) are underlined.

Figure A-6 illustrates the use of both orthographic- and typographic-based devices to promote reading comprehension. Compound words (e.g., *Bundesländer*, *Rechtschreibung*, *Einspruchstermins*) have been divided into their simple components with a red vertical bar (|). Line spacing has been increased significantly. A simple graphic (⌐⎯⎯⌐) is also used to direct the reader's attention to the separable prefix verb *mitteilen*.

RECHTSCHREIB-REFORM NICHT MEHR ZU STOPPEN

Kein Einspruch der Länder eingegangen

dpa Bonn - Die 16 Bundesländer haben endgültig grünes Licht für die Reform der deutschen Rechtschreibung gegeben. Bis Ablauf des offiziellen Einspruchstermins, gestern 12.00 Uhr, wurde kein Veto mehr eingelegt, teilte die zuständige Staatskanzlei Schleswig-Holsteins mit.

Formal muß jetzt noch die Bundesregierung zustimmen, was Bundesinnenminister Manfred Kanther (CDU) bereits signalisiert hat. Der Bund ist für die Schreibweise in den Amtsstuben verantwortlich, die Kultusminister für die Vermittlung der Rechtschreibung in den Schulen. Ein Staatsvertrag mit den deutschsprachigen Nachbarländern Österreich und Schweiz wird voraussichtlich im Juni dieses Jahres die Reform besiegeln. Die neuen Regelungen sollen ab dem 1. August 1998 gelten.

Insgesamt wird die Schreibweise von 185 der 12 000 Wörter des deutschen Grundwortschatzes geändert. Aus 212 Rechtschreibregeln werden 112. Von 57 Kommaregeln bleiben neun übrig, wobei die Interpunktion künftig mehr nach Gefühl als nachstarren Regeln benutzt werden kann. Viele häufig falsch geschriebenen Wörter werden dem gesprochenen Deutsch angepaßt, teilweise auch als Alternativangebot. Grundsätzlich soll eher getrennt als zusammengeschrieben werden, mehr groß als klein.

© DIE WELT, 6.3.1996

Format Features
none

Figure A-1. Source Input Material

RECHTSCHREIB-REFORM NICHT MEHR ZU STOPPEN

Kein Einspruch der Länder eingegangen

dpa Bonn - Die 16 Bundesländer haben endgültig grünes Licht für die Reform der deut-
schen Rechtschreibung gegeben. Bis Ablauf des offiziellen Einspruchstermins, gestern
12.00 Uhr, wurde kein Veto mehr eingelegt, teilte die zuständige Staatskanzlei
Schleswig-Holsteins mit.

Formal muß jetzt noch die Bundesregierung zustimmen, was Bundesinnenminister
Manfred Kanther (CDU) bereits signalisiert hat. Der Bund ist für die Schreibweise in
den Amtsstuben verantwortlich, die Kultusminister für die Vermittlung der
Rechtschreibung in den Schulen. Ein Staatsvertrag mit den deutschsprachigen Nach-
barländern Österreich und Schweiz wird voraussichtlich im Juni dieses Jahres die
Reform besiegeln. Die neuen Regelungen sollen ab dem 1. August 1998 gelten.

Insgesamt wird die Schreibweise von 185 der 12 000 Wörter des deutschen
Grundwortschatzes geändert. Aus 212 Rechtschreibregeln werden 112. Von 57
Kommaregeln bleiben neun übrig, wobei die Interpunktion künftig mehr nach Gefühl
als nachstarren Regeln benutzt werden kann. Viele häufig falsch geschriebenen Wörter
werden dem gesprochenen Deutsch angepaßt, teilweise auch als Alternativangebot.
Grundsätzlich soll eher getrennt als zusammengeschrieben werden, mehr groß als
klein.

© DIE WELT, 6.3.1996

| **Format Features** |
| subject phrases |

Figure A-2. Subject Phrases

RECHTSCHREIB-REFORM NICHT MEHR ZU STOPPEN

Kein Einspruch der Länder eingegangen

dpa Bonn - Die 16 Bundesländer haben endgültig grünes Licht für die Reform der deutschen Rechtschreibung gegeben. Bis Ablauf des offiziellen Einspruchstermins, gestern 12.00 Uhr, wurde kein Veto mehr eingelegt, teilte die zuständige Staatskanzlei Schleswig-Holsteins mit.

Formal muß jetzt noch die Bundesregierung zustimmen, was Bundesinnenminister Manfred Kanther (CDU) bereits signalisiert hat. Der Bund ist für die Schreibweise in den Amtsstuben verantwortlich, die Kultusminister für die Vermittlung der Rechtschreibung in den Schulen. Ein Staatsvertrag mit den deutschsprachigen Nachbarländern Österreich und Schweiz wird voraussichtlich im Juni dieses Jahres die Reform besiegeln. Die neuen Regelungen sollen ab dem 1. August 1998 gelten.

Insgesamt wird die Schreibweise von 185 der 12 000 Wörter des deutschen Grundwortschatzes geändert. Aus 212 Rechtschreibregeln werden 112. Von 57 Kommaregeln bleiben neun übrig, wobei die Interpunktion künftig mehr nach Gefühl als nachstarren Regeln benutzt werden kann. Viele häufig falsch geschriebenen Wörter werden dem gesprochenen Deutsch angepaßt, teilweise auch als Alternativangebot. Grundsätzlich soll eher getrennt als zusammengeschrieben werden, mehr groß als klein.

© DIE WELT, 6.3.1996

Format Features
verbs

Figure A-3. Verbs

RECHTSCHREIB-REFORM NICHT MEHR ZU STOPPEN

Kein Einspruch der Länder eingegangen

dpa Bonn - Die 16 Bundesländer haben endgültig grünes Licht für die Reform der deutschen Rechtschreibung gegeben. Bis Ablauf des offiziellen Einspruchstermins, gestern 12.00 Uhr, wurde kein Veto mehr eingelegt, teilte die zuständige Staatskanzlei Schleswig-Holsteins mit.

Formal muß jetzt noch die Bundesregierung zustimmen, was Bundesinnenminister Manfred Kanther (CDU) bereits signalisiert hat. Der Bund ist für die Schreibweise in den Amtsstuben verantwortlich, die Kultusminister für die Vermittlung der Rechtschreibung in den Schulen. Ein Staatsvertrag mit den deutschsprachigen Nachbarländern Österreich und Schweiz wird voraussichtlich im Juni dieses Jahres die Reform besiegeln. Die neuen Regelungen sollen ab dem 1. August 1998 gelten.

Insgesamt wird die Schreibweise von 185 der 12 000 Wörter des deutschen Grundwortschatzes geändert. Aus 212 Rechtschreibregeln werden 112. Von 57 Kommaregeln bleiben neun übrig, wobei die Interpunktion künftig mehr nach Gefühl als nachstarren Regeln benutzt werden kann. Viele häufig falsch geschriebenen Wörter werden dem gesprochenen Deutsch angepaßt, teilweise auch als Alternativangebot. Grundsätzlich soll eher getrennt als zusammengeschrieben werden, mehr groß als klein.

© DIE WELT, 6.3.1996

Format Features

passive voice
future tense

Figure A-4. Passive Voice and Future Tense

RECHTSCHREIB-REFORM NICHT MEHR ZU STOPPEN

Kein Einspruch der Länder eingegangen

dpa Bonn - Die 16 Bundesländer haben endgültig grünes Licht **für** die Reform der deutschen Rechtschreibung gegeben. Bis Ablauf des offiziellen Einspruchstermins, gestern 12.00 Uhr, wurde kein Veto mehr eingelegt, teilte die zuständige Staatskanzlei Schleswig-Holsteins mit.

Formal muß jetzt noch die Bundesregierung zustimmen, was Bundesinnenminister Manfred Kanther (CDU) bereits signalisiert hat. Der Bund ist **für** die Schreibweise **in den** Amtsstuben verantwortlich, die Kultusminister **für die** Vermittlung der Rechtschreibung **in den** Schulen. Ein Staatsvertrag **mit den** deutschsprachig**en** Nachbarländern Östereich und Schweiz wird voraussichtlich **im** Juni dieses Jahres die Reform besiegeln. Die neuen Regelungen sollen **ab dem** 1. August 1998 gelten.

Insgesamt wird die Schreibweise von 185 der 12 000 Wörter des deutschen Grundwortschatzes geändert. Aus 212 Rechtschreibregeln werden 112. Von 57 Kommaregeln bleiben neun übrig, wobei die Interpunktion künftig mehr nach Gefühl als nachstarren Regeln benutzt werden kann. Viele häufig falsch geschriebenen Wörter werden **dem** gesprochen**en** Deutsch angepaßt, teilweise auch als Alternativangebot. Grundsätzlich soll eher getrennt als zusammengeschrieben werden, mehr groß als klein.

© DIE WELT, 6.3.1996

Format Features

dative case prepositions
accusative case prepositions
2-way prepositions

Figure A-5. Prepositions Governing Dative and Accusative

RECHTSCHREIB-REFORM NICHT MEHR ZU STOPPEN

Kein Einspruch der Länder eingegangen

dpa Bonn - Die 16 Bundes|länder haben endgültig grünes Licht für die Reform der deutschen Recht|schreibung gegeben. Bis Ablauf des offiziellen Einspruchs|termins, gestern 12.00 Uhr, wurde kein Veto mehr eingelegt, teilte die zuständige Staats|kanzlei Schleswig-Holsteins mit.

Formal muß jetzt noch die Bundes|regierung zustimmen, was Bundes|innenminister Manfred Kanther (CDU) bereits signalisiert hat. Der Bund ist für die Schreib|weise in den Amts|stuben verantwortlich, die Kultus|minister für die Vermittlung der Recht|schreibung in den Schulen. Ein Staats|vertrag mit den deutsch|sprachigen Nachbarländern Östereich und Schweiz wird voraussichtlich im Juni dieses Jahres die Reform besiegeln. Die neuen Regelungen sollen ab dem 1. August 1998 gelten.

Insgesamt wird die Schreib|weise von 185 der 12 000 Wörter des deutschen Grund|wort|schatzes geändert. Aus 212 Recht|schreib|regeln werden 112. Von 57 Komma|regeln bleiben neun übrig, wobei die Interpunktion künftig mehr nach Gefühl als nachstarren Regeln benutzt werden kann. Viele häufig falsch geschriebenen Wörter werden dem gesprochenen Deutsch angepaßt, teilweise auch als Alternativ|angebot. Grund|sätzlich soll eher getrennt als zusammen|geschrieben werden, mehr groß als klein.

© DIE WELT, 6.3.1996

> **Format Features**
> compound word division (|)
> line spacing (21pt)
> separable-prefix verbs ⌐⌐

Figure A-6. Compound Words and White Space

Appendix B. Cue Detection in German

Rapid detection and processing of surface-level cues are essential to fluent reading comprehension. The successful foreign language student learns to quickly recognize the importance of these markers and to use them in the reading comprehension process. The complexity of the morphological surface features of a language and the information these features can convey are not to be ignored as can be seen in the following few examples from German.

The importance of case-markers in a strongly case-based language such as German is usually overlooked by the beginning student, particularly if his or her native language is one (such as English) where the case system—if not dead—is moribund. German makes extensive use of case-markers in its definite and indefinite articles, and adjective and noun endings, to convey meaning. *Die bräutliche Schwester befreite der Bruder* (Wagner, *Die Walküre*) means "The brother freed his sister and bride" and not, as word order would suggest, "The bride-sister freed her brother." (The second definite article is the key word: *der* is the nominative form, not the accusative [*den*], of the masculine definite article.) Because there are only 6 distinct forms of the German definite article available to make 24 possible grammatical distinctions (3 genders, 2 numbers, and 4 cases), the correct interpretation in this example is a matter of context and the gender and number of the accompanying noun (hence, the teacher's exhortation to the beginning German student to master the gender when learning vocabulary).

In some instances, purely surface-level cues can be used to decode a construction. Masculine nominative definite descriptions (beginning with *der*) when used with an adjective can be easily distinguished from feminine genitive or dative, or genitive plural forms, by noting the *e* ending on any accompanying adjective. (The feminine genitive or dative, or genitive plural definite article forms (which also use *der*) require an *en* adjective ending.) When used with one (or more) adjectives, the German masculine nominative definite article consists of a reliable cue pattern: d*er _e, ..., _e* that is readily distinguishable from the feminine and the genitive plural patterns: d*er _en, ..., _en*. The proficient reader (and listener) has automatized the recognition of German articles and accompanying adjective and noun endings as a single cue pattern that maps to the appropriate semantic function; the system being proposed here

can help to facilitate this pattern recognition process by allowing the user to globally highlight the pattern and thereby encourage the student to note, look for, or otherwise focus upon its role in the text. The highlighting device itself (color, font style, point size, underlining) is selected by the user and may even be chosen to reflect some mnemonic the student wants to associate with the form to be emphasized.

Case distinctions are only one of many confusing aspects of a foreign language to beginning language students. In German, the extended adjective construction, relative pronouns, "two-way" prepositions (and prepositional phrases, in general), separable verbs, the passive, the subjunctive are all sources of exasperation in the language classroom. But it is easy to see how judicious use of simple formatting tools (coupled with the computer's correct syntactic understanding of these troublesome constructions) can keep the reader from getting bogged down and thus moving further along the input comprehension process.

In each extended adjective construction, it would be possible and helpful to simply italicize the first and last word of the expression. Consider, for example, a line taken from Theodor Fontaine's *Frau Jenny Treibel* and used by Strutz (1981) to illustrate the troublesome extended adjective construction:

> ...*der* anderen, mit Geschmack und Sorglichkeit gekleideten und trotz ihrer hohen Fünfzig noch sehr gut aussehenden *Dame*....

In German, the relative pronoun in relative clause constructions is mandatory. Moreover, the most commonly used form is identical (except in the dative plural and all genitive cases) to the definite article. The relative pronoun construction is widely used, however, and quick recognition of it as a modifier of an antecedent noun (or pronoun) is crucial for understanding authentic German. Word order is helpful: a relative clause is a subordinate clause with the finite verb positioned at the end of the clause. The clause itself is set off by commas. But since the pronoun's case is determined by the role it plays in its own clause, and its gender and number must agree with its antecedent, the only (syntactic) clue it provides as to its referent is its gender and number. We believe it would be useful in acquiring the relative pronoun construction (in German) by underscoring the antecedent-relative pronoun link whenever it appears in the text, with the form of the "underscoring" chosen by the user. Consider, for example, the following:

> Am 5. Mai 1990 fand in Bonn *die erste Runde* der sogenannten "Zwei-Plus-Vier"-Verhandlungen der Aussenminister statt, *die* dann am 11. September 1990 zu *einem Bertragsabschluss* führten, *der* freie Bahn für die Deutsche Einheit schuf. [*Deutschland Nachrichten*, German Information Center, New York, NY, September 1995]

Here it becomes obvious that the relative pronoun, *die*, has as its antecedent *die erste Runde,* and that the second relative pronoun, *der*, refers to *[der] Bertragsabschluss.*

German prepositions fall into three major classes, depending generally on which case they take (or govern): genitive, dative, or accusative. One set of nine very common prepositions, however, takes either the dative or accusative, depending on whether the preposition is intended to indicate location (dative) or destination (accusative). For the student coming to grips with this distinction, a program that systematically underscored (or otherwise highlighted) the distinction would be of value. In the following, for example, different fonts (italics and boldface) are used to underscore the two different uses of two-way prepositions.[2] The two prepositions governing the dative (location), along with the operative verb, are italicized; the two prepositions governing the accusative (destination), again with the operative verb, are in bold face.

> Mit diesen Worten **stellt** er die Tasse, das Glas und den Fisch **auf den Tisch** und
> **geht an einen anderen Tisch**. *An diesem Tisch sitzt* eine junge Amerikanerin.
> Sie liest die Speisekarte und macht ein trauriges Gesicht. Der dicke Hund des
> Restaurants *schläft unter ihrem Stuhl.* [Goedsche and Spann 1994]

Clearly, similar techniques can be used to help German students deal with separable prefix verbs (with the prefix coming at the end of the sentence or clause) and other forms (such as the passive) that requires the student to recognize and process long-distance dependencies. Other ways in which modest typographic capabilities can be used in a language learning environment to promote reading comprehension include regular or on-demand highlighting of the root word used in an adjective, adverb, or compound-noun construction (under the assumption that it may be more easily recognizable in its root form).

---

[2] Different colors or more subtle (and less distracting) differences in font size could also be used.

Appendix C. Market Survey

There are many commercial products available that implement some of the ideas presented in this paper. This appendix lists, briefly describes, and provides contact information for a few prominent representatives in this area. The material is organized into four sections: general purpose foreign language learning software, lexical analyzers, text-to-speech synthesizers, and time-scale modification software. This listing is not intended to be comprehensive.

C.1 General Purpose Foreign language Learning software

There are many CD ROM-based products available in the commercial foreign language education market. The following is a representative list of the major products.

C.1.1 The Rosetta Stone

The Rosetta Stone Language Library is a collection of CD-ROMs for learning languages developed and published by Fairfield Language Technologies. Languages currently available include Spanish, French, German, Dutch, Russian, Chinese, and English as a Second Language (ESL). The program provides a browser that allows students to freely page through both textual, audio, and video material; student voice recording and comparison with a native speaker; dictation, prompting the student to write words, phases, and sentences in the target language and then checking their work; tutorials, where students are methodically guided through material and re-exercised on material where the student had difficulty; and scoring and testing.

Fairfield Language Technologies
122 S. Main Street
Harrisonburg, VA 22801
Phone: (540) 432-6166
Fax: (540) 432-0953
E-mail: info@trstone.com
URL: www.trstone.com

C.1.2   Language Connect and Triple Play Plus

Both Language Connect and Triple Play Plus are available from Syracuse Language, Inc.

Language Connect is an innovative distance learning course that combines multimedia software with one-on-one guidance from a professional instructor via the Internet. At this time Language Connect is only available for Spanish I.

Triple Play Plus is a multimedia language immersion program that stresses listening, reading, and speaking in the target language.

Syracuse Language, Inc.
Syracuse, New York
Phone: (800) 797-5264
URL: www.syrlang.com

C.1.3   Language Now!

LanguageNow! is a multimedia-based foreign language learning product from Transparent Language, Inc. It is touted as supporting all four language skills (reading, writing, listening, and speaking) and provides some of the features of the proposed system described in the body of the paper.

Transparent Language, Inc.
22 Proctor Hill Road
PO Box 575
Hollis, NH 03049-0575
Phone: (603) 465-2230
Fax: (603) 465-2779
URL: www.transparent.com

C.1.4   Dynamic Japanese

Many multimedia-based products are becoming available for Japanese. "All language is in context and sequenced so that the language is acquired as students interact through a variety of learning tasks. Interactive exercises focus on basic vocabulary development and an understanding of basic language structures that aid in Japanese language comprehension. Learning tasks include: intensive listening and repetition, comprehension questions, gap fill-ins, matching kana combinations to spoken words, voice recording and comparison to native Japanese speakers (www.dyned.com/dyned/eng/djmain.htm)."

DynEd International
989 E. Hillsdale Blvd.

Suite 130
Foster City, CA 94404 USA
Phone: (650) 578-8067
Fax: (650) 578-8069
URL: www.dyned.com

C.1.5   Learn to Speak

Learn to Speak titles are available in French, German, Spanish, and Japanese. They focus on the use of everyday conversational needs (e.g., greeting strangers, asking for directions) and use a variety of traditional techniques, including dialogue, vocabulary, quizzes, and voice recording.

The Learning Company
One Athenaeum Street
Cambridge, MA 02142
Phone: (617) 494-5700
Fax: (617) 494-5898
URL: www.mecc.com

C.1.6   Multimedia German, French, Spanish

The various foreign language software from Sofsource, Inc., focus on helping students achieve the five goals of the National Standards in Foreign Language Education: communication, cultures, connections, comparisons, and communities—the so-called five "C's" of foreign language education. The series offers extensive coverage of grammar, vocabulary, and conversations by native speakers, including discussions of travel, business, and culture.

Sofsource, Inc.
P.O. Box 16317
Las Cruces, NM 88004
Phone: (505) 532-0500
URL: www.sofsource.com

C.2   Lexical Analyzers

Several lexical analysis tools are developed and marketed by Lingsoft, Inc. Three of the more prominent products are:

- ENGCG, a constraint grammar parser for English;
- NPtool, a tool for the detection of English noun phrases;
- TWOL, a morphological analyzers for several European languages, including English, German, Finnish, Swedish, Danish, Estonian, Norwegian, and Russian.

Lingsoft, Inc.
Helsinki, Finland
URL: www.lingsoft.fi

C.3 Text-To-Speech

SounText

SounText is a multilingual voice synthesizer for MS-DOS and Microsoft Windows environments. The standard package supports English, French, German, Italian, and Spanish languages using Berkeley Speech Technology. Mandarin Chinese is available as an option.

Fortress Systems, Inc.
50 Airport Parkway
San Jose, CA 95110
Phone: (408) 289-8818

C.4 Time-Scale Modification Software

ETSM

ETSM (Entropic's Time-Scale Modification) is a speech and signal-rate change software product that permits digital playback of audio data at rates faster or slower than the original recording, but without changes to the local periodicity or sample rate. The speech or other signal will maintain its natural quality with only the duration of the playback changed. The product is designed for use in speech research and other applications where detailed analysis of audio waveforms is desired. These applications include speech labeling, speech pathology, psycho-perception, audio forensics, and others. It can also be used for variable-rate playback in multimedia applications and for fitting audio segments to required lengths.

Entropic Research Laboratory, Inc.
400 N Capitol Street, NW
Suite G100
Washington DC 20001
Phone: (202) 547-1420
Fax: (202) 546-6648
URL: www.entropic.com

Appendix D. Translations of German Examples

Page ES-2   German Press Agency Bonn - The 16 states of the German Federal Republic have finally given the green light to the reform of the German spelling system. According to the chancellery of Schleswig-Holstein—the state officially responsible for the reform measure—no "vetos" were received as of noon yesterday—the close of the period for official objections.

page 5   To be sure it looks just like the alley where Rolf and I ended up that time we got lost.

page 27   Therefore the passageway with the door from which we saw Pia leaving that night would also have to be nearby.

page 46   The street on which he lives is very elegant.

page 3   No More Stopping Spelling-Reform
No Objections Raised by the States

German Press Agency Bonn - The 16 states of the German Federal Republic have finally given the green light to the reform of the German spelling system. According to the chancellery of Schleswig-Holstein—the state officially responsible for the reform measure—no "vetos" were received as of noon yesterday—the close of the period for official objections.

The federal government must still formally approve the measure, something that the federal interior minister, Manfred Kanther (Christian Democratic Union), has already signaled his intention to do. The federal government is responsible for spelling policy in the bureaucracy, the cultural minister is responsible for implementing spelling reform in the schools. An international treaty with the german-speaking neighbors of Austria and Switzerland will probably be ratified in June. The nine rules go into effect August 1, 1998.

Altogether the spelling of 185 of the 12,000 words of the basic German vocabulary will be changed. The number of grammar rules will be reduced from 212 to 112. Of the 57 comma rules 9 will remain, so that in the future punctuation will be more a matter of feel than of hard and fast rules. Many frequently misspelled words either will be changed to reflect spoken German or adopted as

permissible alternative forms. Basically there should be fewer long compound words, and more words will be capitalized.

© DIE WELT [The World], March 6, 1996

page 2   ...[of] the other, dressed with taste and care and in spite of her late 50's still very good looking woman....

page 2   On May 5, 1990 the first round of the so-called "Two-Plus-Four" negotiations of the foreign ministers took place. These talks lead on September 11, 1990 to an agreement which created the way to German unity.

page 3   With these words he places the cup, the glass, and the fish on the table and goes to another table. At this table sits a young american lady. She reads the menu and makes a sad face. The fat dog of the restaurant is sleeping under her chair.

References

Barzun, J. 1991. *Begin here: the forgotten conditions of teaching and learning*. Chicago: The University of Chicago Press.

Bates, E. and MacWhinney, B. 1987. "Competition, variation, and language learning," in *Mechanisms of language acquisition*, edited by B. MacWhinney. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Bernstein, J. 1995. "Speech recognition explained." Proceedings of the Computer Assisted Language Instruction Consortium 1995 Annual Symposium - "Computers & Collaborative Learning," June 19-23, 1995.

Danella, U. 1991. *Die Reise nach Venedig*. München: Albrecht Knaus Verlag.

Dennett, D. 1995. *Darwin's dangerous idea: evolution and the meanings of life*. New York: Touchstone.

Delcloque, P. 1995. "Pronunciation matters(!)." Proceedings of the Computer Assisted Language Instruction Consortium 1995 Annual Symposium - "Computers & Collaborative Learning," June 19-23, 1995.

Durrell, M. 1992. *Using German: a guide to contemporary usage*. Cambridge: Cambridge University Press.

Garrett, N. 1987. "A psycholinguistic perspective on grammar and CALL," in *Modern media in foreign language education: theory and implementation*, edited by W. Smith, Lincolnwood, Illinois: National Textbook Company.

Garrett, N., Noblitt, J., and Dominguez, F. 1990. "Computers in foreign language teaching and research: Part 2." *Educom*, Summer 1990.

Goedsche, C. and Spann, M. 1994. *Deutsch für Amerikaner*. New York: American Book Company.

Herron, C. and Seay, I. 1998. *The effect on listening comprehension of using authentic oral texts in the foreign language classroom*. Atlanta, Georgia: Emory University.

Hofstadter, D. 1997. *Le ton beau de Marot: in praise of the music of language*. New York: Basic Books

Hubbard, P. 1995. "Integrating CALL into a listening comprehension course." Proceedings of the Computer Assisted Language Instruction Consortium 1995 Annual Symposium - "Computers & Collaborative Learning," June 19-23, 1995.

Krashen, S. 1982. *Principles and practice in second language acquisition*. Oxford: Pergamon Press.

Levin, L. and Evans, D. 1995. "ALICE-chan: a case study in ICALL theory and practice," in *Intelligent language tutors: theory shaping technology*, edited by V. Holland, J. Kaplan, and M. Sams. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Lyman-Hager, M. 1995. "Encore et Encore: Reviewing Language Structure via Multimedia." Proceedings of the Computer Assisted Language Instruction Consortium 1995 Annual Symposium - "Computers & Collaborative Learning," June 19-23, 1995.

MacWhinney, B. 1987. "The competition model," in *Mechanisms of language acquisition*, edited by B. MacWhinney. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

_____. 1992. "The competition model and foreign language acquisition," in *Intelligent tutoring systems for foreign language learning*, edited by M. Swartz. Berlin: Springer-Verlag.

Merzenich, M., Jenkins, W., Johnson, P., Schreiner, C., Miller, S., and Tallal, P. 1996. "Temporal processing deficits of language-learning impaired children ameliorated by training." *Science*, Vol. 271, 5 January 1996.

Oxford, R. 1995. "Linking theories of learning with intelligent computer-assisted language learning (ICALL)," in *Intelligent language tutors: theory shaping technology*, edited by V. Holland, J. Kaplan, and M. Sams. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Pinker, S. 1994. *The language instinct*. New York: William Morrow and Company, Inc.

Smeaton, A. 1992. "Progress in the application of natural language processing to information retrieval tasks." *The computer journal*, Vol. 35, No. 3, 1992.

Steiner, G. 1992. *After Babel: aspects of language and translation*, second edition. Oxford and New York: Oxford University Press.

Strutz, H. 1981. *1001 pitfalls in German*. Woodbury, NY: Barron's.

Tallal, P., Miller, S., Bedi, G., Byma, G., Wang, X., Nagarajan, S., Schreiner, C., Jenkins, W., and Merzenich, M. 1996. "Language comprehension in language-learning impaired children improved with acoustically modified speech." *Science*, Vol. 271, 5 January 1996.

Wilson, E. 1998. *Consilience: the unity of knowledge*. New York: Alfred A. Knopf.

Acronyms

AI	Artificial Intelligence
ASR	Automatic Speech Recognition
CAI	Computer Assisted Instruction
CD-ROM	Compact Disc-Read Only Memory (or Media)
dB	Decibel
DoD	Department of Defense
ESL	English as a Second Language
FTP	File Transport Protocol
L2	Second Language
LLI	Language Learning Impaired
HTML	Hypertext Markup Language
Hz	Hertz
MT	Machine Translation
NLP	Natural Language Processing
NNV	noun-noun-verb
PDF	portable document format
SLA	Second Language Acquisition
SOV	subject-object-verb
SP	Speech Processing
SVO	subject-verb-object
TSM	Time-Scale Modification
TTS	Text-to-Speech
URL	Uniform Resource Locator

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE March 1999	3. REPORT TYPE AND DATES COVERED Final	
4. TITLE AND SUBTITLE A Multimedia-Based System for Foreign Language Listening and Reading Comprehension		5. FUNDING NUMBERS Central Research Project C5014	
6. AUTHOR(S) Dale E. Lichtblau			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Institute for Defense Analyses (IDA) 1801 N. Beauregard St. Alexandria, VA 22311-1772		8. PERFORMING ORGANIZATION REPORT NUMBER IDA Paper P-3459	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Institute for Defense Analyses (IDA) 1801 N. Beauregard St. Alexandria, VA 22311-1772		10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, unlimited distribution: May 21, 1999.		12b. DISTRIBUTION CODE 2A	

13. ABSTRACT (Maximum 200 words)

This report presents a conceptual and functional design for a multimedia-based system for foreign language listening and reading comprehension. It describes how integrating existing and proven multimedia and artificial intelligence technology, and the World Wide Web may provide an innovative and effective approach to foreign language learning in DoD. The proposed system uses speech and natural language processing technology to phonetically, morphologically, and typographically "enhance" authentic foreign language input. The goal is to extend and strengthen the language student's listening and reading comprehension ability. The functional design of the system is grounded in Second Language Acquisition theory.

14. SUBJECT TERMS Foreign Language Learning, Multimedia, Artificial Intelligence, World Wide Web, Speech Processing, Natural Language Processing, Second Language Acquisition, German Language.			15. NUMBER OF PAGES 98
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL